

Techniques for Inserting Additional Train Paths into Existing Cyclic Timetables

Von der
Fakultät Architektur, Bauingenieurwesen und Umweltwissenschaften
der Technischen Universität Carolo-Wilhelmina
zu Braunschweig

zur Erlangung des Grades einer
Doktoringenieurin (Dr.-Ing.)
genehmigte

Dissertation

von
Yuyan Tan
geboren am 1. Oktober 1985
aus Hunan, China

Eingereicht am 01. Dezember 2014
Disputation am 27. April 2015

Berichterstatter Prof. Dr.-Ing. Jörn Pachl (TU Braunschweig)
Prof. Dr. Anita Schöbel (Uni Göttingen)

Acknowledgements

The defense of this thesis marks the end of my time as a PhD candidate. I have had four great years in the Institut für Eisenbahnwesen und Verkehrssicherung (IfEV) and would like to use these first pages to thank some of the people who have impacted my life as a PhD candidate and as an inhabitant in Germany.

First and foremost, I would like to express my special appreciation and thanks to my advisor Prof. Dr. Jörn Pachl, who was abundantly helpful and offered invaluable assistance, support and guidance. He gives me a lot of freedom to do what I want to do. I would like to thank him for encouraging my research and for allowing me to grow as a research scientist.

I wish to thank Prof. Dr. Schöbel for serving as my committee member. Thank her for her brilliant comments and suggestions. I also express my warm thanks to my project external coach Prof. Dr. Braband, Dr. Schmidtke and Dr. Erhard at Siemens for their support and guidance.

I would especially like to thank Prof. Dr. Nie from Beijing Jiaotong University. She has been a tremendous mentor for me. I can not say thank you enough for her tremendous support and help. I feel motivated and encouraged every time I converse with her. Her advice on both research as well as on my career have been priceless.

I also have to thank Assoc. Prof. Dr. Jiang from Tongji University. Without his knowledge and assistance this study would not have been successful. My thanks and appreciations also go to my colleague in developing the project and friends willingly helped me out with their abilities. Tecker, Birgit, Anne, Kai, Niels, Tobias, Po-Chi, David, Iris, Gunnar, Petra, Friederike and Yiyang support me in adapting to the new life in Germany, and incited me to strive towards my goal. I am grateful for their constant support and help. Not forgetting to my Chinese friends from Beijing Jiaotong University who always been there. I am not lonely when you join me, walk beside me and help me along the way continuously.

Deepest gratitude are also due to China Scholarship Council (CSC) and International Rail Automation Graduate School (iRAGS) of Siemens for the financial support.

Special thanks to my beloved families; for their understanding and endless love, through the duration of my studies. In addition, I wish to express my love and gratitude to my beloved husband, who spent sleepless nights with and was always my support in the moments when there was no one to answer my queries.

Abstract

With the development of high-speed railway (HSR) in the world, cyclic timetable shows several obvious advantages in transport marketing and train operation planning. However, due to too many passenger origins and destinations (ODs) with long distance and some special trains such as night trains, the pure cyclic timetable is not suitable in China's HSR. As a consequence, a hybrid timetable concept named "cyclic + non-cyclic" timetable is possible and proposed in this thesis, which means that only part of trains are scheduled periodically in the cyclic timetable and other trains are scheduled as non-cyclic trains. This pattern of timetable is constructed with a cyclic core timetable in which the trains of long-distance or low-frequency are inserted as extra trains. Nowadays, models and algorithms for cyclic timetabling have been well developed but the technique of inserting additional train in an existing cyclic timetable is still a significant demand for research. Moreover, the technique of inserting new train services also can be applied in short-term planning that concerns the re-construction of a generic timetable in order to adapt to the demands of the individual weeks or days, such as national holidays or major sports events that generally require an increase of train services.

Based on the background facts, this thesis deals with the *adding train paths* (ATP) problem for scheduling additional train services in an initial cyclic timetable. This problem is of considerable difficulty and must be performed in practice.

The ATP problem firstly is an integration of timetable scheduling and rescheduling problem. Train dispatcher both modifies the given timetable to manage the disruptions in existing operations and establishes schedules for extra trains. The ATP problem therefore is considered involving many general constraints, such as flexible running times, dwell times, headways and time windows. Characterized based on an event-activity graph, a general mixed integer program (MIP) model for the ATP problem is formulated. In addition, several extensions to the general model are further proposed. The real-world constraints that concerning the acceleration and deceleration times, priority for overtaking, station capacity, allowed adjustments, periodic structure and frequency of services are incorporated into the general model. In order to get a new timetable that with low deviations to the initial services and high quality of the performance to the additional trains, objective functions of minimizing travel time of additional trains, minimizing total adjustments of initial trains, minimizing the makespan and maximizing the robustness of the new timetable are discussed in this thesis.

More importantly, many additional trains may not be inserted because of a shortage of train-set capacity, which would be a very limited resource when the frequency of train services is high. Consequently, how to cover the entire trains with minimum train-sets must be also taken into account in the ATP problem in order to obtain a match between the requested additional trains and the available number of train-sets. The timetable scheduling, train-set planning and rescheduling are three complex optimization problem respectively and usually solved in a sequential manner. In this thesis, we integrate these phases into the ATP problem in a model that decides simultaneously initial trains' modifications, additional trains' schedules and train-set circulation.

The train-set circulation in the ATP problem is decomposed to two sub-problem. (i) For current train-set circulation, the initial train-set route is assumed to be fixed; it is solved as a rescheduling problem of a tight constraint to keep the current train-set circulation. (ii) For additional trains, the train-set circulation problem is equivalent to covering all the additional trains with minimal number of train-sets. The difficulty of the second sub-problem is that train-sets circulation are usually determined in the tactical planning phase after all of the train lines and timetable have been fixed. However, in the ATP problem, the additional trains do not exist in the initial timetable, and even the number of additional trains depends on the number of instantly available train-set at the right place. In order to solve the problem in a reasonable time, we start from fixed train-set route, and then apply flexible train-set route that provides possible alternative turning activities to decrease the waiting time of a train-set in an overnight turn-around.

Case studies based on Shanghai-Hangzhou HSR line in China investigate the proposed framework and associated techniques. Meanwhile, the performance of various settings are compared to analyse the affecting factors to this specific problem.

Contents

1	Introduction	1
1.1	Developments of High-speed Railway in China	1
1.2	Aim and Relevance of the Thesis	2
1.2.1	Social Relevance	4
1.2.2	Scientific Relevance	4
1.3	Review of the Related Literature	5
1.3.1	Literature Review on Timetable Scheduling and Rescheduling Problem	6
1.3.2	Literature Review on Adding Train Paths Problem	11
1.4	Contributions and Outline of the Thesis	13
1.4.1	Contributions of Work	13
1.4.2	Outline of the Thesis	14
2	Appropriate Timetable Mode for China’s High-speed Railway	16
2.1	Non-cyclic Timetable	16
2.2	Cyclic Timetable	20
2.2.1	Cyclic Timetable Experiences Around the World	20
2.2.2	Factors Affecting the Adaptability of Cyclic Timetable	23
2.3	“Cyclic+non-cyclic” Timetable	27
2.3.1	Applicability of “Cyclic+non-cyclic” Timetable	27
2.3.2	Planning Railway Operation for “Cyclic + non-cyclic” Timetable	29
3	Definition of Adding Train Paths Problem	35
3.1	Model Assumptions	35
3.1.1	Infrastructure Assumptions	35
3.1.2	Train Assumptions	36
3.1.3	Strategies for Inserting Extra Trains	36
3.2	Objectives in Adding Train Paths Problem	37
3.2.1	Deviations to Initial Timetable	38
3.2.2	Travel Time	38
3.2.3	Robustness	39
3.2.4	Makespan	39

3.2.5	Costs	39
3.3	Modeling Requirements	40
3.3.1	Trip Time and Dwell Time Constraints	40
3.3.2	Safety Headways	40
3.3.3	Periodic Constraints	42
3.3.4	Connection Constraints	43
3.3.5	Frequency Constraints	43
3.3.6	Time Window Constraints	44
3.3.7	Station Capacity	44
4	Mathematical Model for Adding Train Paths Problem	45
4.1	A General Adding Train Paths Model	45
4.1.1	Railway Network Input	45
4.1.2	An Event-activity Graph	46
4.1.3	Model and Integer Programming Formulation	49
4.2	Extensions of the Adding Train Paths Problem Model	54
4.2.1	Acceleration and deceleration time	54
4.2.2	Priority for Overtaking	55
4.2.3	Station Capacity	57
4.2.4	Tolerance of Disruptions for Initial Timetable	58
4.2.5	Frequency of Additional Trains	60
4.3	Objectives Functions	61
4.4	Experimental Studies	64
4.4.1	Purpose	64
4.4.2	Scenarios and Experiment Setting	65
4.4.3	Result	67
4.4.4	Summary of the Computational Results	74
5	Integration of Train-set Circulation and Adding Train Paths Problem	80
5.1	Use of Train-set in China's HSR	80
5.1.1	Fixed Use of Train-set	80
5.1.2	Flexible Use of Train-set	81
5.2	Terminology and Literature	82
5.2.1	Terminology	82
5.2.2	Literature Overview	84
5.3	Special Features of Train-set Planning in the Adding Train Paths Problem	85
5.4	Formulation for Train-set Routing in Adding Paths Problem	89
5.4.1	Fixed Train-set Route	93
5.4.2	Flexible Train-set Route	95
5.5	Experimental Studies	101
5.5.1	Test Problem	102

5.5.2	Results	104
5.5.3	Summary of the Computational Results	111
6	Conclusions and Future Work	113
6.1	Main Results	113
6.2	Answering the Research Questions	115
6.3	Limitations of the Thesis and Recommendations for Further Research . .	117
	Appendices	127
A	Different Mode of Maintenance Time	128
B	The timetable in Guangzhou-Shenzhen HSR line	130
C	Graphical User Interfaces of the adding train paths tool	131

List of Tables

1.1	Comparison of timetable scheduling, rescheduling and adding train paths problem	9
1.2	Characteristics of adding paths problem and solution approaches.	11
2.1	HSR network	22
2.2	Arrangement of Tokaido Shinkansen train stopped within an hour (down) (Sun et al. (2011))	24
2.3	Punctuality of passenger transport in Germany	25
2.4	Recommended capacity consumption by UIC 406	34
4.1	General subscripts and parameters	50
4.2	Input parameters for inserting train paths	51
4.3	Decision variables.	51
4.4	CPU times (in seconds) for inserting 10-20 trains	71
4.5	Set of instances with different tolerance of disruptions	71
4.6	Result for the instances of inserting 10 trains	76
5.1	Results from experiments with different tolerance of disruption	105
5.2	Objective value of the cases 14-20 from Table 5.1 with different time limits	108
5.3	Results from experiments using same scenarios as for experiments presented in Table 5.1, but with objective function (1b)	108
5.4	A comparison of inserting additional trains with and without time window	110

List of Figures

1.1	Adding train paths problem	5
1.2	The adding train paths problem model	14
2.1	Different operating time periods between continuous and discontinues timetable	19
2.2	Different operating time periods between trains of different speed	19
2.3	High-speed railway models according to the relationship with conventional services	21
2.4	Trains of long journey time in cyclic timetable	24
2.5	Applicability analysis of “cyclic + non-cyclic” timetable	29
2.6	Information flow between planning phases	30
2.7	The process of choosing cyclic and non-cyclic train ODs	31
2.8	The process of “cyclic + non-cyclic” timetable planning	33
3.1	Principles of assigning headways	41
3.2	Various strategies of rescheduling	43
4.1	Train path P_t and its corresponding activities E_{train}	47
4.2	Headway arcs between two consecutive trains on a same section	48
4.3	A small railway network with three stations and three trains, and the corresponding event-activity graph	49
4.4	Illustration for acceleration time and deceleration time	55
4.5	Train priority at station	56
4.6	Different robustness results from different insertion	63
4.7	Optimal train paths for a request on maximum robustness	64
4.8	One hour time-space diagram for the track between Shanghai and Hangzhou	65
4.9	Overview the values of the performance measure <i>robustness of the insertion</i> based on objective function applied and number of additional trains for each scenario	68
4.10	Overview the values of the performance measure <i>makespan of the new timetable</i> based on objective function applied and number of additional trains for each scenario	68

4.11	Overview the values of the performance measure <i>average trip time of additional trains</i> based on objective function applied and number of additional trains for each scenario	69
4.12	Different insertions to maximise robustness based on different tolerance of frequency constrains $\beta = 10$ or $\beta = 20$	70
4.13	Trends of the objectives based on the initial timetable (5:00-14:00)	72
5.1	Fixed use of train-set	81
5.2	Flexible use of train-set	81
5.3	A train-set route plan	83
5.4	Planning phases beforehand	86
5.5	3 required train-sets in the original timetable	87
5.6	2 required train-sets after adjustment	87
5.7	Planning phases covered by adding paths problem with our contribution	88
5.8	A simple train network	90
5.9	A train cycle	92
5.10	Examples of various insertion solutions and train-set rotations to the same route of $t_1 \rightarrow t_2 \rightarrow t_3 \rightarrow t_4 \rightarrow t_1$	94
5.11	A example with fixed train-set route	95
5.12	The selection of turn around connection (t_2, t_7)	97
5.13	The selection of turn around connections (t_2, t_5) and (t_4, t_7)	97
5.14	Alternative turn around connections on overnight stops	98
5.15	Insertion solution based on flexible train-set route	98
5.16	Time-space diagram for Shanghai-Hangzhou high-speed railway	103
5.17	Speed and stops schedule for additional trains	104
5.18	Solutions with different level of tolerance	107
5.19	Information of additional trains	110
5.20	Experiments using different strategies of train-sets	111
6.1	Turn around activity in a through station	117
A.1	Rectangular-shape maintenance time in conventional rail line	128
A.2	V-shape maintenance time in conventional rail line	128
A.3	Maintenance time in China's HSR line	129
B.1	The timetable in Guangzhou-Shenzhen HSR line	130
C.1	The ATP problem without train-set circulation	131
C.2	The ATP problem with train-set circulation	132
C.3	Show result	132

Chapter 1

Introduction

1.1 Developments of High-speed Railway in China

Railway transportation serves as a back-bone of transportation systems in many countries. In response to the increasing competition with other transportation modes, the railway companies strive to continuously improve the level of service.

Over the past three decades, rail passenger traffic in China has been growing more rapidly than any other railway except India, and far more rapidly than in the E.U. . China's railway transportation as the main mode of public transportation, especially for medium to long distances, it carries more passenger-miles than any other of the worlds railways. China's railways carry over 25% of the world's total passenger-miles, 60% more than the entire E.U. and 136% of Japan's total (Thompson and Tanaka (2011)).

However, the railway in China have been criticized due to the less than optimal usage of infrastructure capacity utilization. In fact, China has an almost unique problem in that, along with its huge population, its rail network carries the highest traffic density, including freight and passenger traffic, in the world and is the critical artery of both freight and passenger travel. As a result, it faces congestion at many points. With the economy grows and increasing demands are placed on the rail sector, more and more certain specific problems are emerging.

In order to solve this problem, China's real solution is to build more rail capacity. One efficient way of doing so is constructing new passenger dedicated lines (PDLs) to (i) relieve the pressure of both passenger and freight demand on its overcrowded existing rail system, (ii) improve transportation connections between the countrys different regions, and (iii) promote the economies of less developed regions (Feigenbaum (2013)).

Faced with the need of new PDLs, China has decided that the new lines should be high-quality and designed to be operated at speeds of 200 to 350 km/h. The governments long-term goal was to make its high-speed rail the largest, most extensive, and most accessible rail network in the world. After years of "speed-up" campaigns, the HSR in China was introduced on April 18, 2007. According to "The Mid and Long-Term

Railway Network Development Plan of China (2008 version)”, by 2020 the total mileage of China’s PDLs will reach 16,000 km and will link all provincial capital cities and cities with a population of over 500,000. With an axial network structure of four main vertical and four main horizontal PDLs, the current travel time will approximately be reduced by 65% from south to the north, and by 75% from east to west (Wang et al. (2012)).

The overall construction and operation of the HSR system has been highly successful. Starting from zero in 2008, China’s operational high speed rails extended to 11,028 km by the end of 2013, including the world’s longest line, the 2,298 km Beijing-Guangzhou HSR line. Currently China is ranked No.1 in the world in terms of its high speed rail mileages, which is more than 50% of the total length of the worlds high-speed lines. Moreover, the planned speed of the fastest services 350 km/h is faster than any service that is currently provided (Thompson and Tanaka (2011)). In addition, the HSR systems daily ridership has grown from 237,000 people in 2008 to 1.33 million people, almost 500 million passengers a year in 2013 (KPMG Global (2013)).

The development of HSR in China is very rapid, however, there are many differences and inadequacies in theory and practice, which brought enormous obstacles for the HSR development and lay hidden for continuous running of the HSR in future. Besides the rare opportunity of development, the China’s HSR is faced with critical challenges from the rapid increasing demand of traffic and higher requirement of organization.

1.2 Aim and Relevance of the Thesis

Timetable is an important component of railway transportation organization. The aim of a train timetable is to define the departure and arrival times for a number of trains on a certain party of railway network. As one of fundamental technical documents, railway timetable provides a basis for synchronizing all scheduling activities over the rail network, and plays an important role in ensuring traffic safety and adapting to market demands. The quality of timetable will directly affects the efficiency of rail transport, which has also been drawing the attention of researches for decades.

Timetable construction is a complex task, in which a trade-off between frequency of train services and infrastructure capacity utilization has to be provided. It aims at determining a periodic timetable for a set of trains that does not violate track capacities and satisfies some operational constraints. The trains have to be run every period of a given time horizon, for example, every day or every hour of a time horizon lasting several months. Roughly speaking, from the view of periodic time horizon, generally there are two main modes for a timetable, named (i) cyclic timetable and (ii) non-cyclic timetable, respectively.

With the development of HSR in the world, cyclic timetable shows several obvious advantages in transport marketing and train operation planning. Since the introduction of cyclic timetables in the Netherlands, many other European countries and Japan have

adopted the concept. After the application for decades, these countries have already accumulated a wealth of experiences in scheduling and applying the cyclic timetable.

However, China is vast in territory and the HSR network covers a wide range with long distance lines and complex structures. For some HSR lines such as being independent with other lines or relative short, the cyclic timetable is also applicable. Whereas, for some lines with complicated situation, it is not suitable to apply a pure cyclic timetable.

The primary research question needs to be solved in the operation management of China's HSR is the following,

What is the appropriate timetable mode for China's high-speed railway?

Recently there are extensive studied over the timetable model including the cyclic timetable mode and non-cyclic timetable mode based on China's HSR, such as Jia (2011), Wang (2008), Xu and Yang (2011), Y et al. (2009) and Li et al. (2011). A hybrid timetable concept named *"cyclic + non-cyclic" timetable* is possible and proposed, which means that only part of trains are scheduled periodically in the cyclic timetable and other trains are scheduled as non-cyclic trains (Nie et al. (2010a), Yang et al. (2010)).

Nowadays, the models and algorithms for cyclic timetabling have been well developed, but the technique of inserting additional trains in an existing cyclic timetable is still a significant demand for research. Consequently,

this paper considers how to alter an existing cyclic train timetable to include additional train services without breaking the initial periodic structure and minimize the number of required train-sets.

To that end, the main research question breaks down into the following sub-questions in this paper:

Why is the "cyclic + non-cyclic" timetable mode more appropriate in China's HSR, and what is the process of planning such a hybrid timetable ?

What real-world requirements should be taken into account, what are the criteria for assessing the quality of an insertion, and how can they be modelled?

What adjustments need to be made to the adding train paths model to integrate the train-set circulation, and how can these adjustments be modelled?

How can the models arising from the previous two questions be solved in a reasonable amount of time?

1.2.1 Social Relevance

The background setting discussed at the start of this chapter directly indicates the social relevance of studying the appropriate mode of timetable. Choosing an optimal mode of timetable or operation has become a key factor for the success of this large infrastructure project of China's HSR. If the HSR wanted to sustain a continuous and healthy development, its operating organization must take into account the convenience of passengers travelling. The "cyclic + non-cyclic" timetable mode, which considers both the characteristics of China's HSR and the advantages of cyclic timetable, will definitely influence the further development of China's HSR services.

As the key issue in scheduling a "cyclic + non-cyclic" timetable, the technique of inserting extra non-cyclic trains in a predefined cyclic time can contribute to the practical tactical planning for a generic timetable. This generic timetable consists of both cyclic trains and non-cyclic trains. It is created by first constructing a pure cyclic timetable, then removing a number of services in off peak hours and finally inserting non-cyclic trains.

Moreover, the technique of inserting new trains also can be applied in short-term planning that concerns the re-development of a generic timetable in order to adapt to the demands of the individual weeks or days, such as national holidays or major sports events that attract a lot of people, that generally require an increase of train services. The additional trains are inserted while taking the structure of the planned timetable into account. This is done to perturb the according existing train services as little as possible or similarly within acceptable levels.

1.2.2 Scientific Relevance

The adding train paths (ATP) problem considers how to rescheduling an existing timetable by scheduling a number of additional trains. This problem is of considerable difficulty and must be performed in practice. Firstly the ATP problem is an integration of scheduling and rescheduling problem. Train dispatcher both modifies the given timetable to manage the disruptions in existing operations and establishes schedules for extra trains.

More importantly, many additional trains may not be inserted because of a shortage of train-set capacity. A *train-set* is the physical unit of rolling stock to cover a train trip, and composed of a set of passenger cars and power units(s). Train-set is a very limited resource when the frequency of train services is very high. Train-sets have to be scheduled to serve the timetable with ever growing demand for capacity, and the railway company must provide the trains with the adequate train-sets. The train-set circulation is usually determined in the tactical planning phase after all of the train lines and timetable have been fixed. However, in the ATP problem, the additional trains do not exist in the initial timetable, and even the number of additional trains depends on the number of instantly available train-sets at the right place. It means, the train-set flow, which imposed by the

timetable trips including scheduled and additional trains, is probably not feasible given the limited available of train-sets. Consequently, how to cover the entire trains with minimum train-sets must be also taken into account in the ATP problem in order to obtain a match between the requested additional trains and the available number of train-sets.

In conclusion, this paper will give an account of how to reconstruct an initial cyclic train schedule by inserting additional train services, see Figure 1.1. From the view of social relevance, this dissertation will contribute to an optimal mode of timetable that corresponds with the actual operation conditions in China's HSR. The ATP problem occurs not only in the tactical planning phase when scheduling a generic "cyclic + non-cyclic" timetable, but also in the short-term planning phase to adapt the increase of passenger flow. The timetables will be convenient passengers' travelling and simultaneously enhance competitiveness of railway companies. From the view of scientific relevance, many authors solve the timetable scheduling, rescheduling and train-set planning independently. However, the three complicated optimization problems of railway system need to be combinationally solved in the ATP problem.

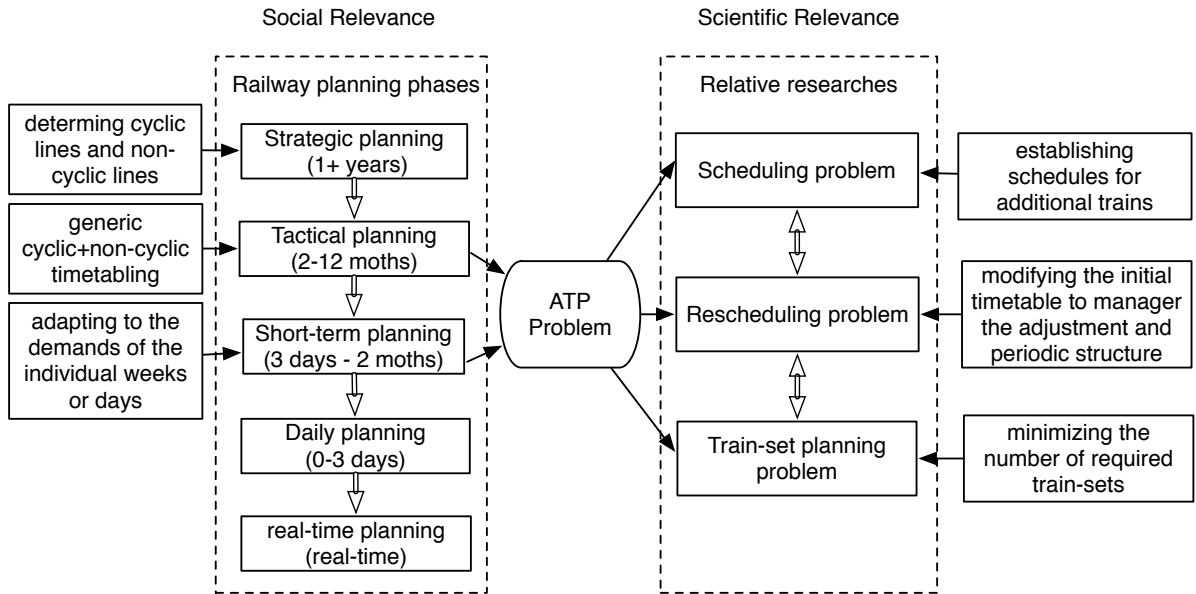


Figure 1.1: Adding train paths problem

1.3 Review of the Related Literature

As analyzed in Section 1.2.2, the ATP problem firstly is an integration of rescheduling and scheduling. There are indications that some of the previous models and techniques could be modified and adapted to solve the ATP problem. This section will provide an overview of the research in railway scheduling, rescheduling and the specific adding train paths problem.

A distinction is here made between the railway timetable scheduling, rescheduling

and the adding trains paths problem. Scheduling (or timetabling) is the process of constructing a schedule from scratch. Rescheduling (or dispatching) indicates that a scheduled timetable already exists and will be modified in case of disturbance or disruption. The specific problem named adding train paths problem in this paper is an integration of rescheduling and scheduling. It constructs a new timetable from an initial timetable with a number of new trains, and simultaneously modifies the existing schedules due to conflicts.

1.3.1 Literature Review on Timetable Scheduling and Rescheduling Problem

In recent years, train timetable scheduling (TTS) and train timetable rescheduling (TTR) problems have a great deal of attentions. For example Törnquist (2006) presents a list of foremost papers published on the area of rail timetable optimization between 1980 and 2006, and a recent survey by Hansen (2009) also summarized emerging methods and solution techniques for train timetabling and dispatching. There are varied models are used to formulate timetable scheduling and rescheduling problem.

Train timetable scheduling problem

The TTS problem has been widely studied in the literatures. From surveys on the problem, it consists of the cyclic and non-cyclic versions. Distinctions of model and algorithm are made between scheduling non-cyclic and cyclic timetables.

Most authors that study cyclic timetabling problem use the models that are based on the Periodic Event Scheduling Problem (PESP), which is introduced by Serafini and Ukovich (1989). Leon (2003) considers a PESP based model for the cyclic railway timetabling problem for NS. His model takes into account the main requirement of Dutch timetables, such as connections, synchronization, variable trip times, rolling stocks, etc.. Lindner (2000) uses the PESP formulation to solve a combination of railway timetabling and railway line planning. His research considers constructing a cost optimal train schedule, which is a timetable that minimizes the cost of the corresponding rolling stock plan.

Several approaches for non-cyclic timetabling have been suggested. They have quite different focus with respect to infrastructure characteristics, objectives and organisation. We refer to Törnquist (2006) and Meng and Zhou (2014) for surveys on this problem. Regarding the objectives applied, no real dominating objective function could be found but there is a tendency towards minimising the total travel time and tardiness. Maximising robustness, profit, and line's frequency are some other examples.

Ghoseiri et al. (2004), Pacciarelli and Pranzo (2001), Zhou and Zhong (2005), Castillo et al. (2011) consider minimising travel time in TTS. Ghoseiri et al. (2004) use Pareto optimality to model the multiple objectives: (i) min fuel consumption, (ii) min passenger

time. An e-constraint method and distance-based method are introduced to determine visiting order on segments and stations allowing meets and overtakes. Zhou and Zhong (2005) considers the generation of a timetable for double track railway applications in China with multiple objectives. Constraints of acceleration and deceleration times are taken into account. By applying two practical priority rules the integer programming model for the second objective criterion can be decomposed and formulated as a number of single train sub-problems which are sequentially solved using branch and bound algorithm. Castillo et al. (2011) applies a MIP model with variable speed, and use sharp upper bounds of the objective function based on the bisection method and reduce the number of binary variables by ignoring those associated with inactive constraints. Based on a MIP model, Zhou and Zhong (2007) adapts three approaches reduce the solution space: (i) a lagrangian relaxation based lower bound rule is used to dualize segment and station entering headway capacity constraints, (ii) an exact lower bound rule is used to estimate the least train delay for resolving the remaining crossing conflicts in a partial schedule, (iii) a tight upper bound is constructed by a beam search heuristic method.

Oliveira and Smith (2001), Mackenzie (2000) and Mu and Dessouky (2011) schedule train timetables with the objective of minimum tardiness. Oliveira and Smith (2001) formulates the problem based on Alternative Graph model, and determine order of trains using B&B and hill climbing. Mackenzie (2000) and Mu and Dessouky (2011) consider flexible train paths in a MIP model. Mackenzie (2000) allocates discrete time units of blocks to trains using (i) lagrangian relaxations, (ii) problem space search local search heuristics. Mu and Dessouky (2011) decomposes the large railway network into smaller sections. Greedy algorithm is used to select route, and genetic algorithm is used to construct the timetable based on the preselected route.

Cacchiani and Toth (2012) reviews the models and algorithms on scheduling an robust timetable. Fischetti et al. (2009) deals with the problem with a MIP model and propose procedure contains two steps: (i) generating an optimal timetable and (ii) finding a robust solution, given fixed event precedences. Shafia et al. (2012) including variable speed in a MIP model, a branch-and-bound algorithm, along with a new heuristic beam search (BS) algorithm is designed to solve the large-size problem. Abbas-Turki et al. (2011) aims to maximize the line's frequency based on a simulation-based model. Rail timetable is treated as a homogeneous traffic and without overtaking. A genetic algorithm to quickly define the sequence of trains for reducing the period is presented.

Train timetable rescheduling problem

In case of disruptions or disturbance occur, the timetable must be rescheduled to resolve the conflicts. Trains may be adjusted by retiming which modifies times for departing or arriving stations, and reordering which select a new route from a set of feasible routes inside or between stations. In order to determine a conflict-free schedule, we refer to Cacchiani et al. (2014) for surveys on this problem, *Alternative Graph Model* and *Event-Activity*

Graph Model are extended and transferred to describe and formulate train operation problem.

Much research related to TTR at a microscopic level is based on the Alternative Graph Model which is introduced by Mascis and Pacciarelli (2002) for no-store job shop scheduling. No-store constraints imply that a job cannot leave a machine until the subsequent machine becomes available. This situation happens in train scheduling as well, which train is viewed as a job and each blocking section is a machine.

D’Ariano et al. (2007b) proposes a fixed speed model and variable speed model for find a conflict-free timetable in real time after train operations are perturbed. Simple dispatching rules, a greedy heuristic based on the alternative graph formulation, and a branch-and-bound algorithm are evaluated in this paper.

D’Ariano et al. (2008b) considers the problem of managing disturbance in real time. In this paper, a real time traffic management system called ROMA (Railway traffic Optimization by means of Alternative graph) is introduced. This problem is decomposed into two sub-problems, one is reordering which is solved by branch and bound algorithm and the other is rerouting which is solved by a local search algorithm. The two sub-problems are then solved iteratively. Other papers focusing on the Alternative Graph Model are following: Corman et al. (2010a), Corman et al. (2009), Corman et al. (2010c), Corman et al. (2010b), Corman et al. (2012), Corman et al. (2011), D’Ariano et al. (2008b), D’Ariano et al. (2007a), D’Ariano et al. (2008a), D’Ariano and Pranzo (2008), D’Ariano et al. (2007b). Different delay scenarios were considered in these papers and all experiments were carried out with ROMA.

A first Mixed Integer Programming (MIP) formulation based on Event-Activity Graph Model is given in Schöbel (2001b) and further developed in Schöbel (2007) and Schachtebeck (2010). The limited capacity of the track system has been taken into account in Schöbel (2009). Schmidt (2013) extend this model by considering rerouting of trains.

Törnquist and Persson (2007) presents an optimisation approach to the problem of rescheduling railway traffic in an n-tracked network when a disturbance has occurred. A MIP model based on graph theory is presented that takes into account reordering and rerouting of trains. For finding a good solution to the problem for large and real-world scenarios, 4 strategies are tested and compared in this paper. Using CPLEX 8.0 it is shown that the solutions with the strategy, that allows certain number of order swaps for specific segments, appears to perform well with respect to computation time and solution quality in many cases. Törnquist (2012) describes a heuristic greedy approach for the same problem in Törnquist and Persson (2007). In order to quickly find a good solution the heuristic performs a depth-first search that branches according to a set of criteria. The heuristic provides solutions for the instances that could not be solved by the approach in Törnquist and Persson (2007) and proved to be good enough.

Acuna-Agost et al. (2010) also presents a MIP model to the TTR problem. They extend the model in Törnquist and Persson (2007) in two aspects. The first is consid-

ering the acceleration and deceleration time; the second one is the modification of some constraints to admit more than one train in the same section running in the same direction. The problem is solved by limiting the search space around the original timetable. Moreover, hard and soft fixing of integer variables with local-branching-type cuts are used. The approach is tested on two different railway network system, thereby the first includes 67 trains in a time horizon of 7 hours, and the second includes 40 trains and a time horizon of 24 hours. Using CPLEX 11.1 it is shown that solutions with an average optimality gap of less than 1% may be obtained in less than 5 min of computation time. Acuna-Agost et al. (2011) studies the same problem and use the same MIP model as in Acuna-Agost et al. (2010). An solution approach called Statistical Analysis of Propagation of Incidents (SAPI) is developed to get a good enough solution in short time. This method proceeds by estimating the probability that an event in the railway network is affected by a set of disturbances, and by reducing the search space accordingly. The method iterating of SAPI + CPLEX find good solutions with an average gap of 0.5% in 128 s in the first network and 0.05% in 15 s in the second network.

Some papers study the problem of delay management from passenger orientation with criteria that minimize the passenger’s delays or deviations of connections, such as Schöbel (2007) decides if connecting trains should wait in a station for delayed trains or departure on time. Schöbel (2009) and Schachtebeck and Schöbel (2010) include constraints on the limited capacity of tracks. A branch and bound algorithm and several heuristic approaches are designed to solve the problem. Dollevoet et al. (2012) extends the problem with rerouting passengers to minimize the total passenger delay. The model is a large MIP model based on Schöbel (2001b) and solved with CPLEX 11.1. Additional constraints are model the routing decisions and the 3 out of 4 cases can be solved within one minute. Based on it, Dollevoet (2013) and Dollevoet et al. (2014) consider constraint of station capacity. An iterative way is used to solve the large size model, with CPLEX 12.2 on an Intel Core i5-2410M with 4 GB RAM, within 3 min.

Comparison

As indicated previously, the approaches and models of TTS and TTR problems could in theory be adapted for the ATP problem too. Above all, the distinguishes between these three problems should be identified in mind. Table 1.1 represents the comparison of timetable scheduling, rescheduling and adding train paths problem.

Table 1.1: Comparison of timetable scheduling, rescheduling and adding train paths problem

Timetable scheduling		Timetable rescheduling	Adding train paths problem
Computation time	Off-line	Real-time	Off-line ¹
requirement			

(Continued on next page)

Table 1.1 – (continued from previous page)

	Timetable scheduling	Timetable rescheduling	Adding train paths problem
Scope of problem	Lines/network	Small dispatch district	Lines/network
Profit	Long-term optimal schedule (months to year)	Real-time planning to solve conflicts	Long-term optimal schedule for generic timetable; short-term planning to satisfy the increased traffic demand
Possible objectives	Min travel time in system (makespan), passenger travel times, waiting times and tradiness; max profit and robustness	Min delay costs, travel time, and passengers' satisfaction	Min adjustments for initial trains, costs, makespan and time window violation; max speed, robustness, and number of the additional trains
Given conditions	Speed, frequency, stop plan (possible bottlenecks, historical performance)	Arrival and departure times at initial timetable	Speed, frequency, stop plan for additional trains; arrival and departure times for initial trains; other limited resource, such as train-sets
Practical decision making process	Limited optimization	Priority based on dispatching rules	Limited adjustments for initial trains; priority based on dispatching rules
Predication horizon	Long-term prediction/forecast	Real-time disturbance detection or propagation	Long-term or short-term prediction/forecast ¹
Traffic situation	Usually assumed ideal	Stochastic disturbances unfold	Usually assumed ideal

¹ The average time requirement of adding freight train paths problem in German railway is about 15 min/train.

Compared with the TTS problem, the main drawback of these approaches in the field of the TTS problem is that they do not explicitly differentiate between initial trains and additional trains. In the ATP problem not all trains have to be added to the schedule, some are already there. It may however be necessary for existing services to be rescheduled in later passes of that approach as changes are forced upon them from the insertion of new services; this needs to be further investigated.

Compared with the TTR problem, since the additional trains could be inserted at any time or within a time window, the ATP problem will lead disruptions to a larger scope in the original timetable and more scheduled trains would be involved. The ATP problem however is not a reactive or dynamic rescheduling problem in the traditional sense since the insertion of additional trains is planned well in advance of when the schedule must be commenced. In addition, in the ATP problem the scheduled trains are allowed to departure either later or earlier than the time in the original timetable, and the violation of cyclic structure of the existing timetable should also be taken into account.

1.3.2 Literature Review on Adding Train Paths Problem

Although the technology of inserting extra trains is very important, there has been few direct related discussion about the ATP problem. The only papers to our knowledge are presented in Table 1.2 which summarizes the studies, like ours, dealing with inserting passenger or freight trains into an exiting timetable.

Table 1.2: Characteristics of adding paths problem and solution approaches.

Publication	Background	Model	Constraints	Objective	Solution	Infrastructure and problem size evaluated for
Burdett and Kozan (2009)	(P)	MIP	(FS), (TW), (SC)	Minimize the total weighted time window violations and the makespan.	(CA), (HA)	(D, S), (N, L), (6/3/48/1-5, 6/5/279/1-5, 10/10/88/1-5, 24/10/127/1-5, 15/5/196/1-5, 54/30/494/1-5, 20/20/202/1-5, 20/12/220/1-5, 20/24/403/1-5)
Tan (2013)	(P)	(MIP)	(FS), (CN)	Minimize consecutive delay	(BB), (AG), (SP), (LS)	(D), (L), (36/60/240/1-15)
Cacchiani et al. (2010)	(F)	(ILP)	(FS)	Maximize the number of additional trains and minimize the violations to the ideal insertion	(HA)	(D,S), (N), (679/202/1440/24-48-96, 520/202/1440/24-48-96, 0/202/1440/188-338-554-64)
Ingolotti et al. (2004)	(F/P)	(CSM)	(FS), (TW), (PC)	Minimize the average traversal time of new train	(DP), (PR)	(S), (N), (81/65/-/20)
Flier et al. (2009)	(F/P)	(LPM), (CSM)	(VS), (TW)	Supports railway planners by computing a set of Pareto optimal solutions with respect to travel time and expected delay to additional trains.	(SP)	(D), (L), (-/-/-/1)
Chapter 4 of this paper	(P)	(MIP)	(VS), (TW), (PC), (PS), (SC), (FC), (CN)	Consider 4 objectives: minimize the total adjustments for initial trains, minimize the average travel time of additional trains, minimize the makespan of the new timetable and maximize the robustness of the new timetable	(BB)	(D), (L), (79/9/420/10-20, 159/9/1440/10-20)

(Continued on next page)

Table 1.2 – (continued from previous page)

Publica- tion	Back- ground	Model	Cons- traint	Objective	Solu- tion	Infrastructure and problem size evaluated for
Chapter 5 of this paper	(P)	(MIP)	(VS), (TW), (PC), (PS), (TS)	Minimize the total ad- justments for initial trains and minimize the number of required train-sets	(BB)	(D), (L), (79/9/420/10-12-14-16)

Symbol descriptions:

- Background: Passenger trains insertion (P), Freight trains insertion (F).
- Model: Mixed integer programming (MIP), Computer simulation model (CSM), Integer linear programming (ILP), Linear regression model (LPM).
- Constraint: Fixed speed (FS), Variable speed (VS), Time window (TW), Connection (CN), Station Capacity (SC), Priority constraint (PC), Periodic structure (PS), Train-set or rolling stock (TS), Frequency constraint (FC)
- Solution: Constructive algorithm (CA), Alternative graphs (AG), Shortest Path Algorithm (SP), Branch-and-bound (BB), Heuristics algorithm (HA), Dynamic programming (DP), Local search (LS), Practical rules (PR).
- Infrastructure and problem size evaluated for: (1) symbol in the first parenthesis: Double-track (D), Single-track (S); (2) symbol in the second parenthesis: Network (N), Line (L); (3) symbol in the third parenthesis: represent problem size, number of initial trains / number of stations or block sections / tested time horizon (min) / number of additional trains. Symbol '-' between double / means missing the information.

Burdett and Kozan (2009) and Tan (2013) solve the problem of inserting passenger trains based on an Alternative Graph Model. Burdett and Kozan (2009) proposes a inserting process that consists of 3 phases by fixing or unfixing some scheduled services. The station capacity is modelled as a capacitated intermediate storage area (buffer) implicitly. The buffer occupancy violations are identified and resolved. Constructive algorithms and improved meta-heuristic are applied to minimize makespan. Tan (2013) inserts additional trains in real-time. In order to meet the limited time requirement and minimize deviations to the existing timetable, it is not necessary to take all of the scheduled trains into considerations. The modification of trains which usually consists of removing or reordering is implemented if and only if it potentially leads to a better solution. The ATP problem is decomposed into two sub-problems in this paper. One is finding the optimal insertion for a fixed order timetable and the other is reordering trains. The two sub-problems are solved iteratively until no improvement is possible within a time limit of computation.

Cacchiani et al. (2010) and Ingolotti et al. (2004) solve the problem of inserting freights trains with assumption that all of the initial trains can not be changed. Cacchiani et al. (2010) inserts a large number of trains that with predefined ideal departure and arrival time. Meanwhile, alternative routes are taken into account when conflicts occur and minimum stopping time at each station that must visit is satisfied. In Ingolotti et al. (2004),

an iteration method is adopted to the ATP problem. Additional trains are inserted at a randomly fixed time belonging to the time window at each iteration and priority rule is predefined for each overtaking and meeting, then the satisfaction of constraints where these are involved is verified. When a constraint is not satisfied, a guided backtracking is done. The technique reduces the search space allowing us to solve real and complex problems efficiently. Flier et al. (2009) computes a set of Pareto optimal train schedules with respect to risk and travel time. Their method aims at finding robust train paths in the sense that the additional train has a low risk of delay upon arrival at its final station and supporting railway planners by computing a set of recommended train paths for a given train request.

This paper is different from the previous ones in two things,

- In Chapter 4, we model the ATP problem with several additional real-world constraints, such as the frequency constraint, the robustness of insertion and the tolerance of adjustments, especially the violation of periodic structure to the initial cyclic timetable. They are considered in light of the practical concerns. Firstly, the deviations to initial timetable are limited, otherwise it will turn to a timetabling problem which has been extensively studied in previous literatures. Moreover, the periodic structure is essential to provide convenience services and consequently should be kept to the most possible extent in case of disruptions. Secondly, the frequency constraint guarantees the regular train services instead of concentrated distributions. Thirdly, the robustness of insertion constrains the extra trains are allocated in the period where the capacity is lightly utilised.
- Furthermore, we consider bicriteria objectives in Chapter 5, minimizing the total adjustments and simultaneously minimizing the number of required train-sets, to the ATP problem. The difficulty in this problem is that train-sets circulation are usually determined in the tactical planning phase *after* all of the train lines and timetable have been fixed. However, in the ATP problem, the additional trains do *not* exist in the initial timetable, and even the number of additional trains depends on the number of instantly available train-set at the right place.

1.4 Contributions and Outline of the Thesis

1.4.1 Contributions of Work

- (1) Based on the background of HSR in China, this paper proposes a hybrid timetable concept named “cyclic + non-cyclic” timetable, with a cyclic core timetable, in which trains that can not be periodically scheduled are inserted as extra non-cyclic train paths. By analysing the characteristics of China’s HSR and summarizing the

experiences in countries in which cyclic timetable is adopted, the applicability of this hybrid timetable in China's HSR is illustrated.

- (2) A MIP model are built for the ATP problem with the several additional real-world constraints. Various objectives are considered and tested to get appropriate inserting solution in different circumstances. Figure 1.2 shows the considered constraints and objectives in the ATP problem.

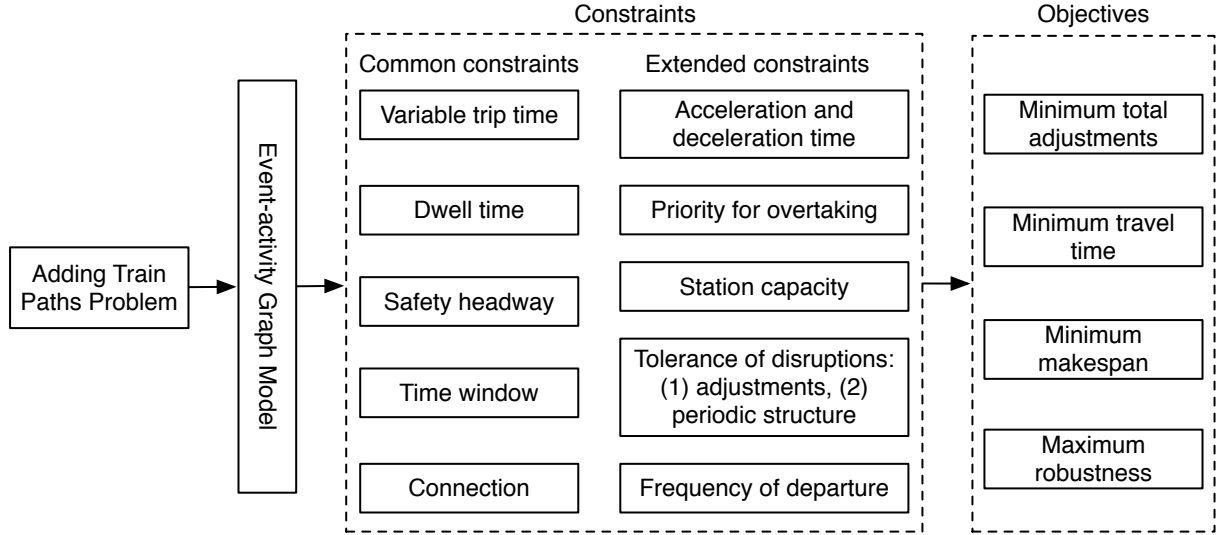


Figure 1.2: The adding train paths problem model

- (3) Train-set circulation is integrated to the ATP problem in order to decide simultaneously initial trains' modifications, additional trains' schedules and train-set routing. The difficulty of this problem is that train-sets circulation are usually determined in the tactical planning phase *after* all of the train lines and timetable have been fixed. However, in the ATP problem, the additional trains do *not* exist in the initial timetable, and even the number of additional trains depends on the number of instantly available train-set at the right place.

Focusing on an existing cyclic timetable on a high-speed passenger rail line, the problem is to minimize both (i) the total adjustments for initial trains and (ii) the number of required train-sets.

- (4) A helpful tool for the ATP problem is developed based on the proposed models and approaches for testing the insertion scenarios.

1.4.2 Outline of the Thesis

The remainder of the thesis is structured as follows. Chapter 2 proposes an appropriate mode of timetable for China's HSR. (i) compares the differences of traffic organization between China's HSR and conventional railway systems in order to explain why the

non-cyclic timetable mode adopted currently is not suitable, (ii) summarizes the cyclic timetable experiences around the world and analyses the factors that affect the adaptability of the cyclic timetable to answer why the pure cyclic timetable mode is not suitable in China's HSR, (iii) a hybrid cyclic timetable, named "cyclic+non-cyclic" timetable, is introduced. By charactering the passenger flow and the requirement of timetable in China's HSR, the applicability of "cyclic+non-cyclic" timetable is illustrated, and (ix) The technological process of "cyclic+non-cyclic" timetabling is also described briefly in this chapter.

According to Chapter 3 which states our assumptions, requirements and the considered objectives, in Chapter 4 we present a mathematical model for the ATP problem. The real-world constraints of general safety, flexible speed, acceleration and deceleration time, allowable adjustment, periodic structure, station capacity, time window, priority for overtaking, and frequency of services are taken into account. Based on a mixed integer programming model, we consider the following objectives: (i) minimizing travel time for additional trains, (ii) minimizing the total adjustment for initial timetable, (iii) minimizing the makespan, (ix) maximizing timetable robustness. A helpful adding train paths tool is developed. Case studies based on Shanghai-Hangzhou HSR line in China consist of two parts. With the fixed initial timetable, the first part investigates the influence of using different objective functions and tolerance of frequency on the values of identified performance measures, while the second part analyses if and how various tolerance of disruptions to initial timetable influences the insertion effect.

Chapter 5 integrates the train-set circulation into the model of Chapter 4. The application methods of train-set adopted in China's CHR are introduced. Focusing on an existing cyclic timetable on a HSR line, the problem is to minimize both (i) the total adjustments for initial trains and (ii) the number of required train-sets. A relaxation approach to the ATP problem is presented. Case studies based on Shanghai-Hangzhou HSR line in China illustrates the methodology and compares the performance of various settings of perturbation tolerance, time window and train-set applications.

Chapter 6 concludes this work with a summary and an outlook to further work. Possible extensions of this thesis are also presented in this chapter.

Chapter 2

Appropriate Timetable Mode for China's High-speed Railway

This chapter explores three different modes of timetable, cyclic timetable, non-cyclic timetable and a hybrid timetable named “cyclic + non-cyclic” timetable. It so gives an impression of the practical environment in which the various railway timetable mode is adopted, of its input factors and operation conditions and of the implications of a timetable for the railway system as a whole.

Section 2.1 considers the non-cyclic timetable which is well applied in China's conventional railway systems. It compares the differences of traffic organization between China's HSR and conventional railway systems, such as the targets of timetabling, scheduling procedure and optimization strategies, maintenance time and operating time for trains, to explain why the current non-cyclic timetable is not suitable for China's HSR. Next, Section 2.2 summarizes the experience of applying cyclic timetable around the world. It analyses the input factors that affect the adaptability of a cyclic timetable from the aspects of operation conditions and passenger flow. By comparing with these characteristics in China's HSR, the pure cyclic timetable is inappropriate. Finally, Section 2.3 introduces a hybrid timetable name “cyclic + non-cyclic” timetable, with a cyclic core, in which the non-cyclic trains are inserted. By charactering the passenger flow and the requirements of timetable in China's HSR, the applicability of “cyclic+non-cyclic” timetable is illustrated. The corresponding process of “cyclic+non-cyclic” timetabling is also described briefly in this section.

2.1 Non-cyclic Timetable

Non-cyclic timetables are generally suited for long distance corridors with heavy dense traffic with limited infrastructure capacity, which is also the only pattern applied in China's conventional railway system. In such a competitive environment, the non-cyclic pattern of the timetables becomes more appropriate than the cyclic pattern, since trains

are asked by so many train operators to be scheduled according to their preferred time, it will become harder to respect a fixed periodic time in a cyclic timetable and simultaneously satisfy most of the requests.

For decades, experts and scholars in China have worked on the potential capacity exploitation and have accumulated a wealth of experiences in non-cyclic timetabling for the China's conventional railway.

However, it is far from being suitable to directly apply these current timetabling models and theories in the HSR lines due to the differences between the conventional railway and HSR system, which are summarized as follows,

Targets of timetabling

In conventional railway, the railway capacity is far from meeting the demands of transport over a long period of time. Therefore, the timetable is scheduled focus on increasing the utilization of capacity. How to reduce time intervals between train operations and make full use of the capacity of the existing infrastructure to increase the frequency of train services are the main targets of timetabling in China's conventional railway.

Whereas, in HSR the timetable should be scheduled with the practical purpose of maximizing the convenience for passengers in order to improve quality of the train services. Not only the reasonable arrival and departure time should be coordinated with public transportation in cities, but also the transfer requirements shall be met especially for cross-line trains running cross at least two different lines. Moreover, maximizing the robustness and minimizing the number of required train-sets are another important goals due to the reliability of a passenger timetable and the expensive costs of train-sets, respectively.

Scheduling procedure and optimization strategies

The traffic in China's conventional railway is organized in a mixed mode that trains on the lines are composed of passenger and freight trains, and there are many different train types of various speed and priority. Since the passenger trains have higher priority than that of freight train, in practice planners construct the timetable by scheduling intercity passenger trains firstly, regional passenger trains, then intercity freight trains and at last regional freight trains. In the process of optimization, passenger trains enjoy a much higher level of priority, freight train has to be overtaken by passenger train whenever conflict occurs.

Meanwhile, at present trains operated on the China HSR are all passenger trains including high-speed trains and medium-speed trains. Typically, a planner first schedules long-distance on-line and cross-line trains corresponding to running within a single line and cross at least two different lines, respectively. medium/short-distance on-line and cross-line trains next, and finally check if the departure and arrival time are designed

in reasonable time, and transfer connections are satisfied. If not, adjustments to the train diagram are executed to meet these operational requirements. Although the trains of various speed and scheduled sequentially, these passenger trains should share nearly equal importance. In HSR, all of the passenger trains have relatively high requirements on departure and arrival times, train travel speed, as well as punctuality rate. It implies that trains shall be of equal priority in timetable optimization or rescheduling in case of disturbance.

Maintenance time

The maintenance of railway infrastructures in China applies a *comprehensive maintenance* mode daily and during which no train is operated. In China’s conventional railway, the comprehensive maintenance time is relatively short and flexible. Under different transport organizational requirements, different mode of maintenance time can be selected, such as “Rectangular-shape type” and “V-shape type”, which are shown in Appendix A.

In China’s HSR, the maintenance standard is at a very high level, which is completely responsible for the technical status of fixed equipments with good conditions and guaranteeing the “safe, comfortable and on time” operation of train services. The time of comprehensive maintenance in HSR needs to have a sufficiently “wide” time interval within the cycle time, typically in the night from 0:00 to 4:00, during which no train is running. This leads to a discontinuous operating time period in timetable. An example is shown in Appendix A.

Operating time for trains

In order to coordinate with urban transport and passenger travel habits, the earliest departure times of passenger trains shall not be earlier than 6 o’clock and the arrival times should not be later than 0 o’clock. In addition, as previous introduced, there is a maintenance time at least 4 hours in HSR, during which no train is running. Due to the discontinuous operating time, the reasonable time window for departures and arrivals in HSR are complexed to calculate. It is not only should be coordinated with urban transport, but also the speed of each type of train. Moreover, outbound and inbound time for cross-line trains should also be taken into account.

Considerations for reasonable departure and arrival times of passenger trains are relatively simple in conventional railway. Generally it only needs to satisfy the requirements to arrange departure and arrival between 6 to 24 o’clock. Although there also exists a off-peak time horizon from 0 to 6 o’clock on the timetable, this period of time can be utilized at a higher degree. As the passenger trains on the conventional line travel at a low speed and with long time, this time horizon can be used for night trains, and moreover, freight trains are operated in this period.

Figure 2.1 shows the different operating time periods between continuous timetable in

conventional lines and discontinuous timetable in HSR. The difference in operating time periods between trains of different speed in HSR is represented in Figure 2.2.

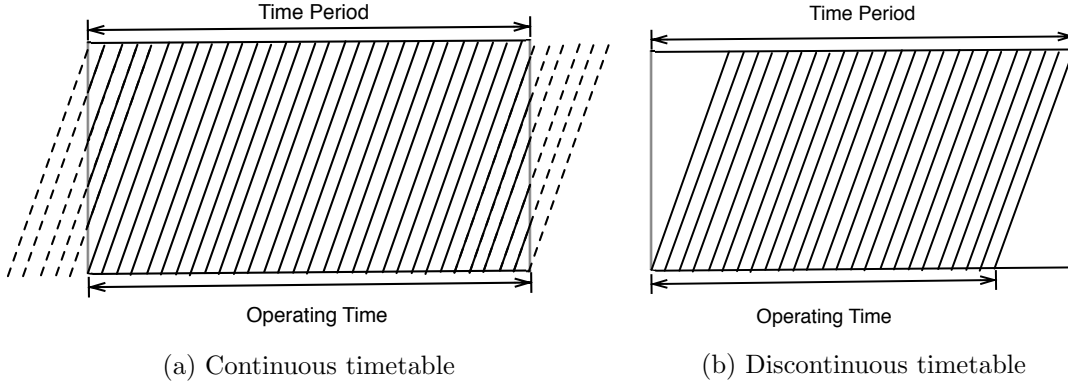


Figure 2.1: Different operating time periods between continuous and discontinues timetable

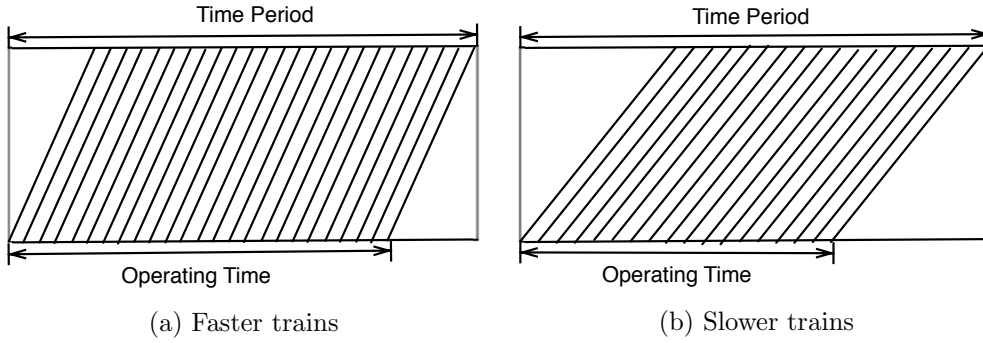


Figure 2.2: Different operating time periods between trains of different speed

Based on the comparisons above, it can be found that although there are fewer train types on HSR than that on conventional railway, the scheduling problem in HSR are much more difficult than that in the conventional lines. It will not work to apply the non-cyclic mode of conventional railway in HSR directly. Moreover, the HSR provide conditions for operating passenger trains that are characterized with high speed, high density and high quality of service. If we follow up current mode of transportation organization in conventional lines, the superiority and efficiency of the high-speed railways will be prevented from giving full play. For example, although the non-cyclic timetable is fun-tuned to the demand of transportation, such a timetable typically contains many trains during the rush hours, and few trains in off-peak hours. Further, a non-cyclic timetable also tends to offer quite different density of traffic in weekend and weekdays. Extremely fine-tuning a timetable to market demand may result in a complex timetable that is hard to consult and to memorize, and also causes difficulties to the organizations for rail companies. Moreover, a non-cyclic timetable concept requires the scheduling of a

timetable for the entire day, for each day of the week. This will typically result in very large planning problems.

2.2 Cyclic Timetable

Since the introduction of cyclic timetables in the Netherlands in 1931, many other European countries and Japan have adopted the concept, especially for the passenger trains' timetables. Sometimes these timetables are also known as “periodic” or “clock-faced ” timetables. Nowadays, cyclic timetable is operated in Germany, Switzerland, Denmark, Austria, Belgium, Great Britain, Norway, Switzerland, Japan and so on. Moreover, many bus and metro systems operate cyclic timetables, though often with smaller cycle times than railway systems do.

In a cyclic timetable, a train for a certain destination leaves a certain station at the same time every cycle time, say every one hour usually. It means after each hour, same pattern of train traffic will repeat it self. Cyclic timetable has several obvious advantages in transport marketing and train operation planning, which are summarized in Leon (2003), shown as follows,

- (1) from the view of customers, train itself and their connections are operated regularly with respect to a cycle time. There is no need for passengers to memorize complex timetables for their regular connection. For example, passengers only have to keep in mind the minutes on which a certain train heads for a certain direction. Moreover, cyclic timetable provides high frequency train services with the same operation time, stop schedule, and transfer connections in each hour. Even an connection is missed due to disturbance, passengers only need to wait at most one hour to transfer to another train, which offers a convenient, comfortable way to travel.
- (2) from a planning point of view, a railway operator can focus the planning on a single cyclic period. In the actual application, the basis of a full-day timetable can obtain by copying the cyclic timetable of one hour for all relevant hours of a day. Clearly, the basis full-day timetable still needs some adjustment, for example, for rush hour traffic or late evening traffic.

2.2.1 Cyclic Timetable Experiences Around the World

From the experiences of Japan and several European countries, the cyclic timetable is proved to be successful. However, it should be noted that the applicability of cyclic timetable is closely related to the operating conditions of the HSR in these countries.

To begin with, we identify the relationship of HSR with existing conventional services and the way in which the use of infrastructure is organized. It is clearly summarized in Campos and de Rus (2009). There are four different exploitation models can be identified, as shown in Figure 2.3 (Campos and de Rus (2009)),

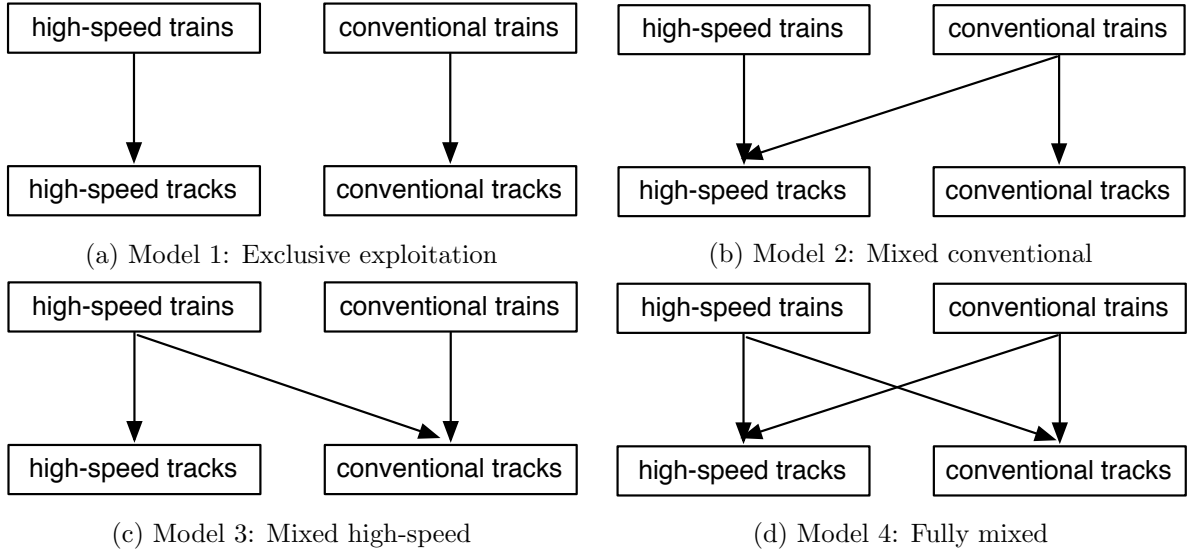


Figure 2.3: High-speed railway models according to the relationship with conventional services

- (1) The *exclusive exploitation model* is characterized by a complete separation between high-speed and conventional services, each one with its own infrastructure, such as in Japan due to the different size of gauge in these two railway system. The important advantage of this model is that the HSR and conventional services and their market organization of fully independent.
- (2) The *mixed conventional model*, where some conventional trains run on high-speed lines, which has been adopted in Spain. This saves the costs of the rolling stock acquisition and their maintenance, furthermore, provides medium-speed trains on certain routes flexibility.
- (3) The *mixed high-speed model* implies that the high-speed trains run both on specifically built high-speed lines, and on upgraded sections of conventional lines, such as in French. One of the major advantage is that this model reduces building costs.
- (4) The *fully mixed model* allows both conventional trains and high-speed trains run on any type of infrastructure, such as adopted in Germany (ICE) and the Rome-Florence line in Italy. The maintenance costs of the infrastructure for this wider use significant increase.

The China's HSR, currently adopts the exclusive exploitation model. However, because of the significant demand of the night trains and the existence of maintenance time, there is a problem that how to coordinate operating of the night high-speed trains and setting up the maintenance time on HSR lines. Some researches suggest cross-line trains, which implies a mixed high-speed model that during the maintenance time the high-speed night trains run off the HSR lines and on the operation of conventional railway lines (Nie et al. (2010b)).

China, France, Germany, Italy, Japan and Spain have six of the most extensive high-speed rail systems in the world (Feigenbaum (2013)). Figure 2.1 summarized the characteristics of the HSR networks in these countries. From this table, we can analyse the factors affecting the adaptability of cyclic timetable, which will be discussed in Section 2.2.2 in detail.

Table 2.1: HSR network

Country	HSR mileage (km)	HSR Model	Passenger data in HSR	Train types and speed	Journey time or distance	Timetable mode
Japan	1,966	Exclusive exploitation; most of trains are running within the region of indepented companies; transfer connections are common between regions	Daily trains: 336; Daily passengers: 409 thousand; Yearly passengers: 149 million	All passenger trains with maximum speed of 270 km/h and only differ in stop schedules	Nozomi: 2.4 h; Hikari: 3 h; Kodama: 4 h	Cyclic
Germany	876	Fully mixed; timetables in both HSR and conventional lines are cyclic; passenger and freight traffic are mixed, and operated at day time and at night time, respectively; transfer connections are common	Yearly Long-distance transport: 131 million	Maximum speed: ICE-1/2: 280 km/h; ICE-3: 330 km/h; ICE-T: 230 km/h	Usually within 4 h	Cyclic
France	1,617	Mixed high speed	Yearly passengers: < 200 million	TGV: 300/320 km/h; THALYS: 300/320 km/h; AVE: 300 km/h	Relative short	Cyclic
Spain	2,326	Mixed convertional	Yearly passengers: 453 million	AVE: 300 km/h	Mostly within 4 h	Cyclic
China	12,625	Exclusive exploitation (currently)	Yearly passengers: almost 500 million	Medium-speed trains: 250 km/h; High-speed trains: 350 km/h	328 passenger ODs with distance more than 2,000 km and 500 passenger ODs with travelling time more than 7 h	Non-cyclic

¹ Mileage Source: International Union of Railways

² Speed Source: HSR Database

³ Passenger Data Source: (1) Japan: take Central Japan Railway Company for example, source: Central Japan Railway Company (2013), (2) Germany: DB 2013 annual report, (3) France: SNCF 2013 annual report, (4) China: KPMG Global (2013)

2.2.2 Factors Affecting the Adaptability of Cyclic Timetable

From the Table 2.1, we can see that cyclic timetable is already adopted in many countries and proved to be efficient. However, not all of the railway systems are suitable for this mode. This section will summarize the factors affecting the adaptability of cyclic timetable based on Table 2.1, which can be considered from the aspects of *operation condition* and *passenger flow*.

Operation condition

The operation condition here mainly refers to the distance of train operation, train types, traffic control system and other facilities.

Distance of train operation

The distance of a train operation should not be too long in a cyclic timetable, as a result of both marketing and capacity utilization. In Japan, France, Spain and Germany, the trip times of most trains are usually within 4 hours. The transfer connections are extremely common between regions to provide long trip for passengers conveniently. However, in China due to the travel behaviour that the passengers usually travel with many luggages and the high population density, the non-transfer trains are preferred, and in practice transfers are seldom even in HSR.

Considering China's vast land area and long distances between major cities, the long distance of operation will lead to long journey times. The statistics in Nie et al. (2010b) show that in China's HSR network, there are 328 passenger origins and destinations (ODs) with distance more than 2,000 km and 500 passenger ODs with travelling time more than 7 h in more than 1,000 passenger flow ODs between major cities, including the regional, provincial capital, and planning independent cities. For instance, in the timetable 2014 of China, the trip times of Beijing-Guangzhou, Beijing-Guilin and Haerbin-Shanghai are 10 h, 12 h and 12 h 30 min, respectively.

Whereas, as analysed in Section 2.1, the maintenance time for China HSR is set from 0 to 4 o'clock, resulting in a discontinuous vision to the timetable. This requires that the journey time for a train should not be too long, otherwise, few cycles of the long-distance trains can be copied in the full-day timetable and consequently will lead to a waste of capacity during the other time cycles in a cyclic timetable, see the shadow areas in Figure 2.4. In fact, the capacity of the shadow areas is difficult to be utilized in a cyclic timetable.

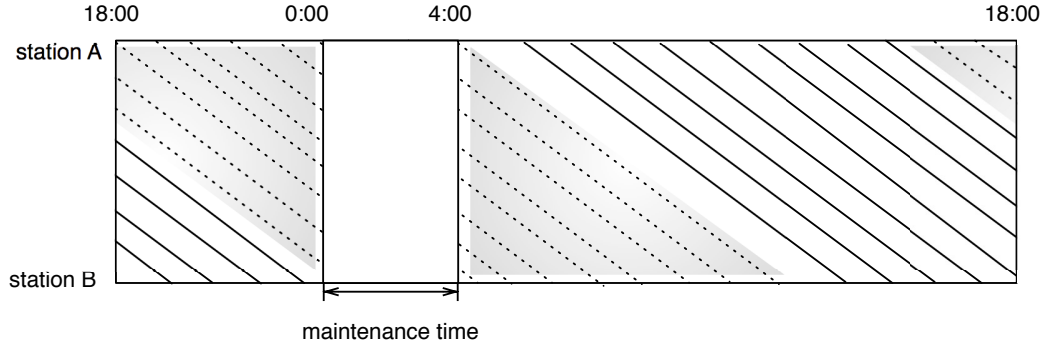


Figure 2.4: Trains of long journey time in cyclic timetable

Train types

Cyclic timetable requires small speed difference between various train types. The existence of large speed difference will significantly cause a decrease of capacity utilization, which will be extended to the entire train diagram with the expansion of the cycle.

The Tokaido Shinkansen in Japan adopts the cyclic timetable with 3 types of train, Nozonmi, Hikari and Kodama. All of the passenger trains have the same maximum speed of 270 km/h and only differ in stop schedules, as shown in Table 2.2 (Sun et al. (2011)).

Table 2.2: Arrangement of Tokaido Shinkansen train stopped within an hour (down) (Sun et al. (2011))

	A		B		A		A		A		C		A		B				A		A		A		C	
Species	N	1	2	3	N	◆	◆	◆					1	2	3	4	◆	◆		N						
Starting time	0	3	3	3	10	13	20	23	26	30	33	33	33	33	33	40	47	50	56							
Tokyo	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Shinagawa	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Shin-Yokohama	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Odawara									●			●														●
Atami									●																	●
Mishima				●					●																	●
Shinfuji									●																	●
Shizuoka		●	●	●					●						●	●										●
Kakegawa									●																	●
Hamamatsu		●	●	●					●								●									●
Toyohashi									●				●			●										●
Mikawaanjo									●																	●
Nagoya	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Gifuhashima									●			●	●	●												●
Maibara									●			●	●	●												●
Kyoto	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Shin-Osaka	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●

Note: (1) A, B, C represent for Nozomi, Hikari, Kodama train; (2) N represent for N700 series train, ◆ represent for temporary trains; (3) • represent for parking, The unlabeled expressed that does not stop passes; (4) 1, 2, 3, 4 under B represent for different stop program in different period at the same

time.

The small speed difference in Japan allows a more intensive usage of high-speed railway infrastructure. In 2013, the Central Japan Railway Company (JR Central) reported that the total daily number of trains is 336 and a maximum of ten Nozomi services per hour; daily passenger capacity of the Shinkansen between Tokyo and Osaka is approximately 320 thousand, which exceeds that of airlines with approximately 29 thousand (Central Japan Railway Company (2013)). In the case of German railway, although a mixed traffic of passenger trains and freight trains is operated, the freight services use the spare capacity of high-speed lines during the night. In addition,

In China, there operate two different trains with maximum speed of 250 km/h and 350 km/h respectively, which also increases the difficulty of cyclic timetabling.

Traffic control system and other facilities

In a cyclic timetable, trains are operated with high frequency and high density. The fundamental facilities, such as high-speed track, high-speed trains and safety infrastructures, with high reliability and intelligentized control systems are necessary to guarantee the punctuality and safety of the train service in HSR.

The Shinkansen employs an ATC (Automatic Train Control) system, eliminating the need for trackside signals. It uses a comprehensive system of Automatic Train Protection. The centralized traffic control manages all train operations, and all tasks relating to train movement, track, station and schedule are networked and computerized. With the new advancements in materials and signaling systems, such as ERTMS/ETCS level 2 and level 3, adopted in Germany, the operators also plan for high-frequency passenger trains on the major corridors.

The timetables in Shinkansen and German railway are very reliable due to the advanced systems. In 2013, JR Central reported that the Shinkansen's average delay from schedule per train was 30 seconds (Central Japan Railway Company (2013)), and Table 2.3 represents the punctuality of passenger transport in Germany.

Table 2.3: Punctuality of passenger transport in Germany

Punctuality passenger transport rail [%]	2013	2012	2011
DB Bahn Long-Distance	73.9	79.1	80.0
Rate of people making connections (long-distance transport/long-distance transport)	86.3	88.7	89.9

* Source: Deutsche Bahn - 2013 Annual Report

Passenger flow

The volume and other characteristics of passenger flow will also influence the adaptability of cyclic timetable.

Volume of passenger flow

The significant characteristics of the cyclic timetable are high service frequency and regular operation. If the volume of passenger flow is small or distributed imbalancedly within a day, the advantages of the cyclic timetable will not be well developed, instead, it will cause insufficient capacity for peak hours and wasted capacity for off-peak hours. For example, the Shinkansens in Japan, counting for 9% of the total operating kilometres of railway, has about 20% of the annual railway passenger ridership (Central Japan Railway Company (2013)), which provides the prerequisite for applying cyclic timetable.

In China, the HSR network by 2020 will cover an area wherein over 90% of the population live and link all provincial capital cities and cities with a population of over 500,000. The large demand of traffic creates the basis condition for cyclic timetable.

Number of passenger ODs

The cyclic timetable is not suitable for the railroad with excessive passenger ODs. Since the stopping plan in a cyclic timetable is relatively fixed, a large number of passenger ODs in a rail network may be adjusted for cyclic scheduling, and thus some passengers may need to transfer since some trains with low service frequency may be modified or possibly cancelled in the adjustment.

Taking Beijing-Tianjin intercity railway in China for example, as the whole line is relatively short with the length of 117 km and passenger ODs is relatively simple, timetables with cyclic patterns have been adopted in this line. Only three stop plans, South Beijing-Tianjin with a average interval of 20 min, South Beijing-Wuqing-Tianjin with a average interval of 50 min and South Beijing-Tianjin-Tanggu with average interval of 120 min are scheduled. This cyclic model with small number of ODs achieved great success.

However, for some long distance HSR lines such as Beijing-Shanghai HSR, at least 222 passenger ODs should be considered on this line (Nie et al. (2010a)). Consequently, in order to adopt the cyclic timetable, one of the key issues is to change the operation mode of “low frequency, long distance and excessive number of ODs” to the mode of “high frequency, medium-long distance, appropriate number of ODs”.

Structure of passenger flow

The structure of passenger flow here mainly refers to the ratio of on-line passenger flow to cross-line passenger flow. In order to guarantee the connections, it must be taken into consideration of the coordination between adjacent lines when scheduling the cross-line trains, which increasing the difficulty of cyclic timetabling. Moreover, a large amount of long distance travel demand brings about a special demand for night service. Actually, on conventional railway network a lot of night trains are operated. The sunset-departure and sunrise-arrival train has become a very successful kind of railway transport product in China. However, during 0:00-4:00 a time window without train services is often arranged

for maintenance work, which will cause great conflict with the operation of night train (Nie et al. (2010a)). In order to solve the conflict, the night trains on HSR line are planned to run on adjacent conventional lines during the maintenance time. Timetable in conventional lines are non-cyclic, the cross-lines consequentially can not be scheduled periodically in the train diagram neither.

It should be noted that the high ratio of cross-line passenger flow does not determine a poor adaptation of the cyclic timetable, which is the comprehensive result of several factors and may be advanced by improving other operation conditions. Taking European HSR for example, the trip time of most trains are relative short that usually within 4 hours, and most of the conventional lines have adopted cyclic timetables. Consequently, although large amount of cross-line trains are operated, the cyclic timetable is still appropriate.

2.3 “Cyclic+non-cyclic” Timetable

After dozens of years of application, both Europe and Japan have formed a set of relatively systematic cyclic timetabling theory.

In China, some of HSR lines have qualification for cyclic timetable, such as Beijing-Tianjin HSR line with the length of 117 km and Guangzhou-Shenzhen HSR line with the length of 147 km, which are featured by single type of trains and the passenger flows are majored by commuter passengers. At present, the timetables of these two lines are scheduled *partly* on a cyclic vision, noting that it is still not a pure cyclic timetable, and have achieved good operating results. Appendix B shows the incomplete cyclic timetable in Guangzhou-Shenzhen HSR line.

2.3.1 Applicability of “Cyclic+non-cyclic” Timetable

However, due to the complex operation condition of some China’s HSR lines, it is not suitable to apply the cyclic timetable. Based on Section 2.2, the reasons mainly be observed and summarized in the following aspects,

- (1) Due to the long-time trip and the discontinuous operating time in the timetable, the trains of long-distance could not have high frequencies. If these trains are scheduled in the cyclic timetable, it will lead to a waste of capacity.
- (2) With the demand of night trains, cross-line trains between HSR lines and adjacent conventional lines are on scheduling. The timetable in conventional railway is non-cyclic, consequently this kind of cross-line trains are difficult to be scheduled periodically.
- (3) The distribution of passenger flow is non-equilibrium both on space and time. A number of low frequency trains are necessary to meet the demand of passenger flow.

However, if these trains of low frequency are scheduled as cyclic trains, it will lead to a waste of capacity. In addition, too many passenger ODs also increases the difficulty to schedule a cyclic timetable.

Furthermore, along with significant advantages, the cyclic timetable may yield higher costs than non-cyclic ones on the other hand. The occupation degree of the late evening trains is much less than during the rest of the day, but a cyclic timetable offers the same train service. Typically, the only way to reduce systems' capacity during off-peak hours is to modify the lengths of the trains. It is believed that off-peak reduction impacts the train-set costs and crew costs, since shorter trains require fewer conductors (Khan (2008)).

In fact, the above mixed operation policy with various types of passenger trains in China's HSR calls for more sophisticated timetable planning methodologies and techniques. Therefore, the cyclic timetable may be considered only in the HSR lines that these inappropriate operation requirements and disadvantages are not prominent. With too many ODs and some special trains, however a hybrid timetable concept named "cyclic + non-cyclic" timetable is possible, with a cyclic core timetable, in which non-cyclic trains are inserted as extra trains.

The "cyclic + non-cyclic" timetable means that only part of trains are scheduled as cyclic trains, such as high-frequency and short-distance trains, and others of low-frequency, long-distance, night trains and cross-line trains are scheduled as non-cyclic trains.

Figure 2.5 indicates the applicability of "cyclic + non-cyclic" timetable in China's HSR.

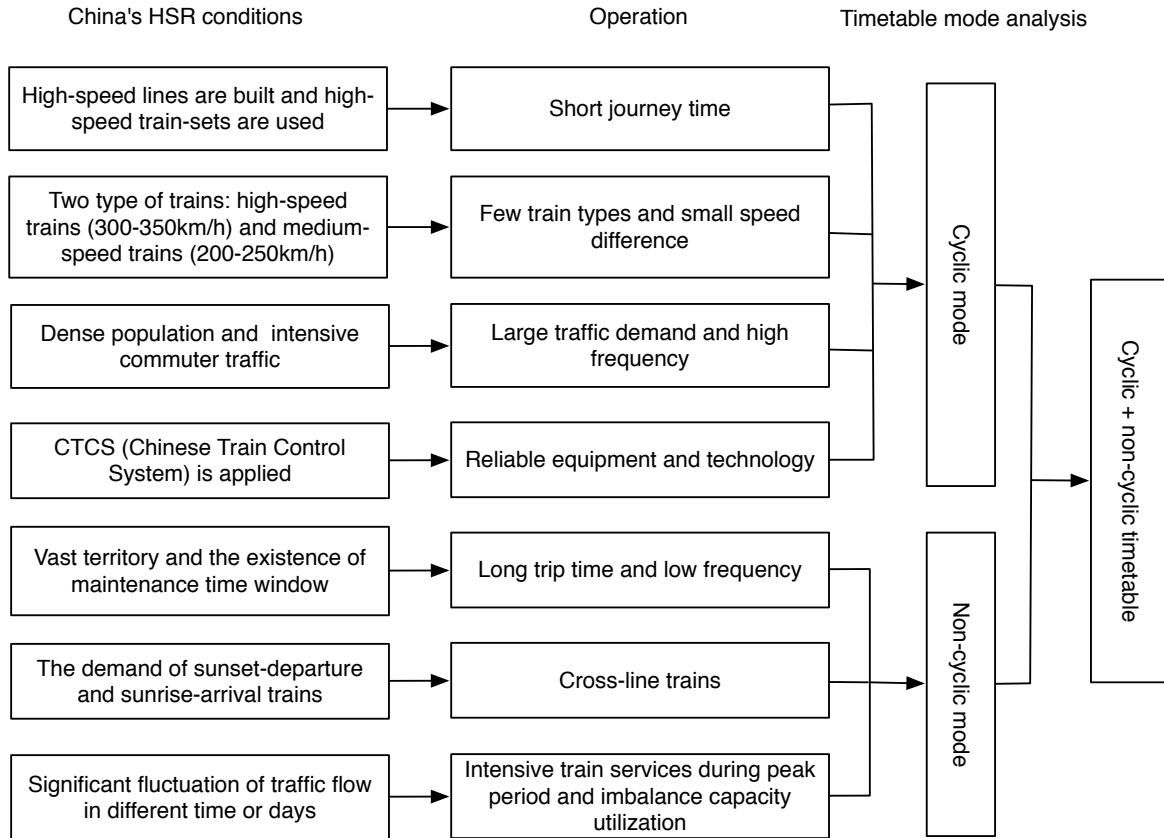


Figure 2.5: Applicability analysis of “cyclic + non-cyclic” timetable

2.3.2 Planning Railway Operation for “Cyclic + non-cyclic” Timetable

The planning of railway operations mainly concerns timetable and the two main resources, the train-sets and crew. The planning of these resources undergoes several phases before the actual operations. This section gives a short description of the procedure of “cyclic + non-cyclic” timetabling and the resources problems in each phases.

The planning phases can be classified by the time horizon of the involved decisions. Nielsen (2011) divide the planning process into five steps depending on the horizon of the decisions, *strategic*, *tactical*, *short-term*, *daily* and *real-time* planning. Figure 2.6 illustrates information flow between planning phases (Maróti (2006)). Rectangles are central planning tasks, ovals are local planning tasks. Arrows indicate the information flow. Dashed arrows represent possible but undesirable influence.

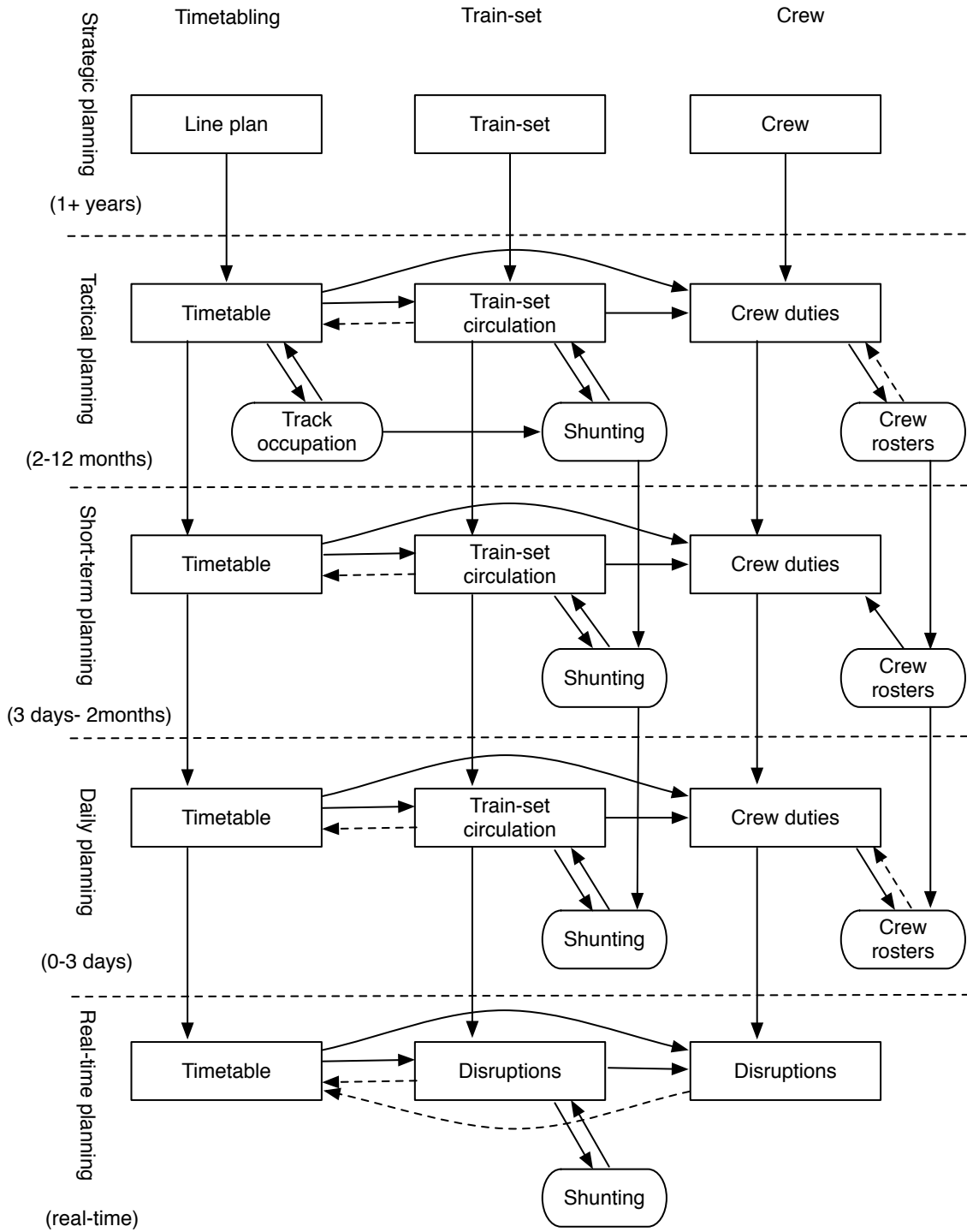


Figure 2.6: Information flow between planning phases

Strategic planning

The horizon of the strategic planning is several years and includes defining the overall objectives of the operator, purchasing and disassembling of rolling stock, hiring and training new crew, the basic structure of the timetable, and decisions on line planning (Nielsen (2011)).

Timetabling is a complicated issue needing take into consideration numerous factors.

At the very first point of timetabling process, the ODs matrix is determined, which specifies the estimated number of passengers on a day between each pair of stations. Based on these forecasts, the train lines are determined. A *train line* is a series of trains that directly connect given stations. This includes decisions on stopping patterns and frequencies of the involved train services.

For “cyclic + non-cyclic” timetabling, there are two key issues. Firstly, “low frequency, long distance and excessive number of ODs” should be optimized to “high frequency, medium-long distance, appropriate number of ODs”. Then the other is to determine which train lines should be scheduled as cyclic trains, and which train lines will be scheduled as non-cyclic trains. The line planning will provide the ODs with n trains per hour as well as that with 1 train per m hours. If the cycle period T is 1 hour, it is clearly that the train train line with only 1 train per day should not be classified into the *cyclic train set*. Otherwise, 23 trains must be removed when extend the one cycle period timetable into full-day timetable. However, if the train lines with 1 train per 2 hours is classified into the *non-cyclic train set*, it will be hard to guarantee the connections for every 2 hours at transfer points. In order to assure the cyclic connection plan among trains and provide convenient service for passengers, it is a trial to plan more trains in cyclic train set. However, in order to control the solution scale of cyclic timetable and decrease waste of capacity, it is better to schedule the trains with low frequency as non-cyclic trains. Figure 2.7 shows the process of identifying cyclic and non-cyclic train ODs (Nie et al. (2010a)).

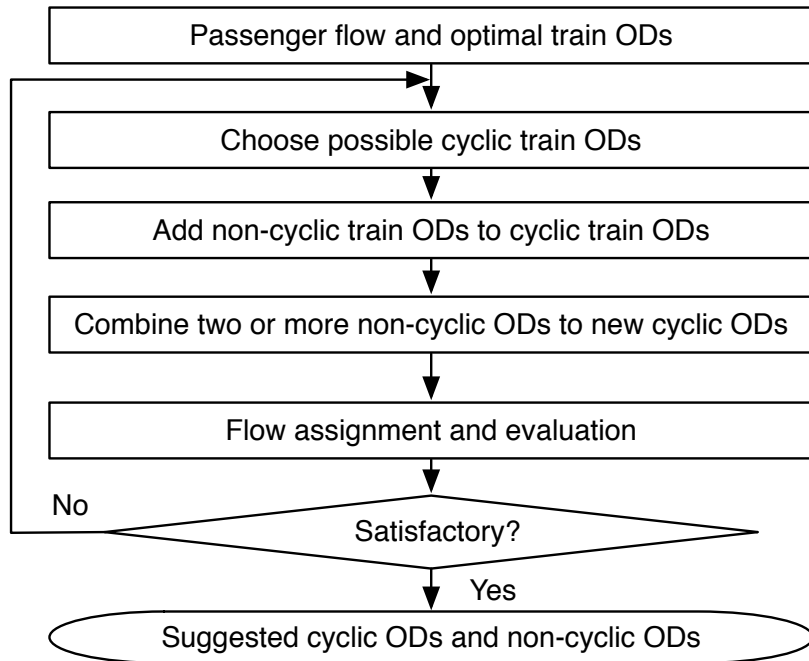


Figure 2.7: The process of choosing cyclic and non-cyclic train ODs

Tactical planning

Tactical planning refers to a planning horizon ranging from 2 months to 1 year. The planning steps conducted during this planning phase include constructing a generic timetable that satisfies service demands, and allocating rolling stock and crew to the generic timetable.

The “cyclic + non-cyclic” timetable is a hybrid pattern with a cyclic core timetable in which the non-cyclic trains are inserted. It is created by first generating a basic one cycle period (usually hourly) timetable.

A crucial extra requirement, then, is that the situation at the end of the cycle period matches the situation at the start of the period. If the start and end situations match, the hourly plan is then copied for all relevant hours of the day to obtain a full-day timetable. This full-day timetable is an cyclic vision in which the departure times are the same for each train connection and for each cycle period. As previously analysed, the departure and arrival times of trains should respect to the maintenance time setting, by removing or adjusting the trains which runs during the maintenance time to match its setting. In practical, a cyclic timetable usually takes market demand into account to a certain extent. Typically, high capacity trains are used during peak hours, that is, 16-cars train-sets are used in China HSR. And in off-peak hours, trains may consist of less carriages, 8-cars train-sets in China HSR. Moreover, from the timetabling view, it is possible that extra trains are operated during rush hours, and some trains are removed during low-traffic hours to obtain a generic full-day cyclic timetable.

In addition, non-cyclic trains are also scheduled in this planning phase. The trains that of low frequency and long distance, and interline trains are inserted to the existing cyclic timetable as non-cyclic trains to form a generic full-day “cyclic + non-cyclic” timetable. Figure 2.8 shows the process of “cyclic + non-cyclic” timetable planning. To obtain this hybrid timetable, the key technologies are the cyclic timetable scheduling and the additional trains insertion, The latter is the research problem in this paper, that is, how to insert non-cyclic trains into the existing cyclic timetable without breaking the initial schedules and the periodic structure.

At the local level, plans are constructed for movements of trains inside the railway stations according to the generic plans, i.e. track allocations based on the timetable, tactical shunting plans based on the train-sets circulation and the crew rosters based on crew duty plan. These local plans primarily serve as a feasibility check for the generic timetable and rolling stock circulation.

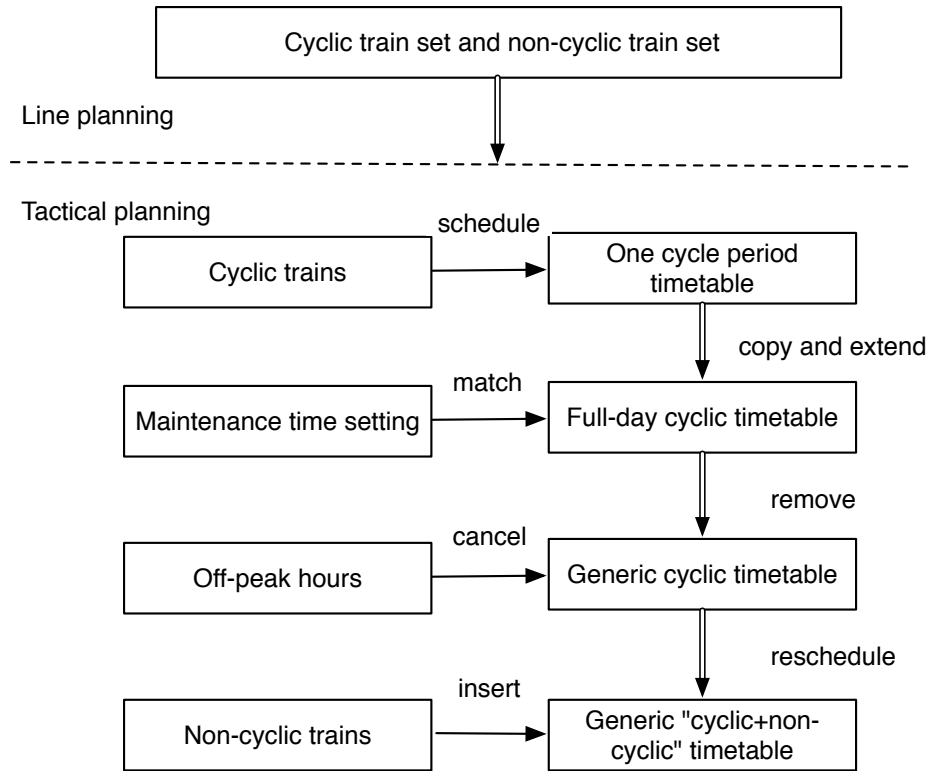


Figure 2.8: The process of “cyclic + non-cyclic” timetable planning

Short-term planning

The short-term planning refers to a planning horizon ranging from few days up to two months. In this phase, the generic plan (i.e. the result of tactical planning) is adapted to the specific demands of the individual weeks or days. Reasons for such adjustments can be the result of the unavailability of tracks (such as maintenance) and the need for extra trains because of increased passenger flow.

For a particular period or unexpected natural disaster and accidents, the unavailability of tracks may require large-scale adjustments and require adaptations of the timetable. Departure and arrival times are adjusted depending on the available infrastructure. The modification usually consists in removing one or a subset of trains from the schedule and re-inserting them back in a hopefully better way. Such adaptations to the timetables also imply changes to the resource schedules. The resources are re-planned and this is done to perturb the generic shunting and crew schedules as little as possible.

Because of cultural or sports events, national feasts, etc., it generally attract a lot of people and require a increased train services. Additional trains are inserted in this planning phases to meet the demand of traffic. The generic timetable and resource schedules are adapted to take this into account. Moreover, the number of additional trains should not only consist with the passenger demand, but also with transportation resource, such as train-set, crew and capacity. Table 2.4 shows the recommended capacity consumption by UIC 406.

Table 2.4: Recommended capacity consumption by UIC 406

Type of line	Peak hour	Daily period	Comment
Dedicated suburban passenger traffic	85%	70%	The possibility to cancel some services in case of delays allows for high level of capacity utilisation.
Dedicated high-speed line	75%	60%	—
Mixed traffic lines	75%	60%	Can be higher when number of trains is low ($< 5/\text{hour}$) with strong heterogeneity.

Daily planning

Daily planning is the last planning phase before the actual operations and has a horizon of up to just a few days. It deals with issues that arise on a daily basis very close to operations. As introduced in (Nielsen (2011)), in this transition some local issues may arise that require minor adaptations. This can be due to temporary unavailability of staff because of illness, or due to unexpected limitations in shunting capacity or in rolling stock availability. Most of these conflicts can be handled locally by exchanging duties between staff or rolling stock, and thus require little or no global coordination. And another important issue is this phase is the preventive maintenance of rolling stock.

Real-time planning

This phase concerns the dispatching of resources during the actual operations. The most important task is delay and disruption management. In real-time planning problems, there is not much time for computations to get the optimal solutions. Mostly, any feasible solution which keeps the railway system moving is good. For rolling stock circulation, during the dispatching phase, there is no time either to come up with completely new schedules. Instead, small and localised adjustments are applied.

Chapter 3

Definition of Adding Train Paths Problem

This chapter describes the adding train paths (ATP) problem in this thesis that inserts additional non-cyclic trains into an existing cyclic timetable.

Section 3.1 describes the basic model assumptions. We distinguish between assumptions with respect to the level of detail of the railway infrastructure, the level of detail of trains, and the considered strategies for inserting extra trains. Next, Section 3.2 describes the objectives that are taken into account. Section 3.3 introduces the requirements in the ATP problem.

3.1 Model Assumptions

In modelling the ATP problem in China's HSR, we assume the following to be given a priori:

- the infrastructure layout of the railway network,
- the trains to be scheduled in China's HSR, and
- strategies for inserting extra trains.

The following subsections elaborate on each of these three assumptions.

3.1.1 Infrastructure Assumptions

The railway infrastructure is considered as a network consist of stations and tracks.

Nodes Each node in the network represents a station in this paper, where trains from several directions arrive and dwell for some time, where passengers transfer between trains, where train-sets may turn before starting their return trip, etc. (Leon (2003)).

Sections Each section is denoted as a collection of one or multiple tracks between two stations, with no intermediate station in between. When multiple parallel tracks

exist between a pair of nodes, each train is assumed to be assigned a priori to one of the available tracks. In this work, for simplicity, we focus our attention on the case that each track is used for one direction of travel which is typically in China's HSR. In addition, a feasible timetable must satisfy the following *track capacity constraints*:

- a minimum time interval between two consecutive arrivals on the same track on the same direction;
- a minimum time interval between two consecutive departures on the same track on the same direction;
- overtaking along a track is not allowed.

The track capacity constraints impose that overtaking between trains occurs only within stations, junction yards, or on side tracks. To the end, a train is allowed to stop at any intermediate station to give the possibilities to some other trains to overtake it.

3.1.2 Train Assumptions

Train is considered in the form of so-called *train line*. which is a direct train connection between an origin and destination station along a given route. Each train line has a frequency specifying how many trains of that line are operated each cycle time. Associated with each train line is a type, determining at which stations the trains of the line call and the velocity of trains.

From the viewpoint of speed, the most common current train types in China's HSR are high-speed and medium-speed trains, which is operated with 300-350 km/h and 200-250 km/h, respectively. For each train, the velocity on sections is assumed to be known a priori. In this paper, we extend the model by incorporating running time supplement to a model with variable trip time.

From the viewpoint of operating pattern, trains consist of cyclic trains and non-cyclic trains. As previously discussed in Chapter 2, the trains of short-distance and high frequency are suitable to be scheduled in a cyclic timetable. The trains such as long-distance, low frequency and cross-line trains should be scheduled as non-cyclic trains and then need to be inserted to the cyclic timetable.

3.1.3 Strategies for Inserting Extra Trains

Before inserting new trains, an important consideration is the utilisation level of the existing timetable and the time frame. The utilisation level of the initial timetable is fairly well known. For example, one can tell whether the railway infrastructure is lightly or heavily utilised by counting the number of scheduled trains and comparing this number with some measure of capacity. Periods of idle times can be used to consider the extent of the utilisation for inserting extra train services.

With these issues in mind, in order to minimize the deviations to the initial timetable, the following strategies are possible for solving the ATP problem,

- **Fixed strategy:** Fix all previously scheduled services.
- **Unfixed strategy:** The initial services can be rescheduled.

When track infrastructure utilisation is light then fixed strategy would be applied. The idle time which usually consists of buffer paths is used for additional trains.

Adding a large number of trains to a timetable that heavily utilises the system or at incorrect periods of time will ultimately be unsuccessful if all of the initial trains are fixed. A relaxation of certain conditions (and tolerances) will be necessary in practice thus promoting some adjustments of scheduled trains. Here, we can borrow the ideas of adjustment strategies from the dispatching or rescheduling problem. Then possible methods for conflicts resolutions in ATP problem are,

- (1) use the train running time supplement to extend or compress the trip time to assure the minimal headway, maintain the periodicity of the initial cyclic timetable, and achieve the purpose of decreased adjustments.
- (2) reduce the stopping time appropriately by strengthening the station operation and management and organizing rapid service, meanwhile, extension of a scheduled stop and additional stops for operational requirement are also taken into consideration.

3.2 Objectives in Adding Train Paths Problem

Several objectives come to mind when inserting additional trains in an existing cyclic timetable. These can basically be divided into four groups:

- small deviations to initial timetable,
- satisfying customers,
- creating a stable and robust new timetable, and
- controlling costs.

Note that these objectives may be conflicting. As an example, passengers of the additional trains would be satisfied if they are offered a short time of travel that without any intermediate stops, however, such a operation would clearly result in large deviations to the existing services.

3.2.1 Deviations to Initial Timetable

An important aspect of the ATP problem is the competition amongst new trains with existing trains for railway infrastructure. The scheduling of the additional train paths must be performed that the disruptions to existing timetable are minimised or similarly kept within acceptable levels. If this requirement is not satisfied then a new timetable that greatly differs from the initial could be obtained. This is a valid though different problem again which has already been significantly addressed as the timetabling problem in previous work.

However, what an acceptable level of disruption is fairly subjective. It should also be mentioned that in some circumstances some train services must be strictly fixed and can suffer no disruption. In addition, the level of acceptable disruption also widely differs according to the train service type. Cross-line trains for example are usually subject to very strict timing. This timing is necessary because of the route of these services take (i.e. usually through both heavily utilised China's HSR and conventional railway networks) and the modifications will propagate to the conventional railway network. Similarly the tolerance of other passenger trains in HSR to delays and alterations is also quite limited.

Furthermore, in order to keep the advantages of the initial cyclic timetable, such as the cyclic behaviour of the train arrivals and departures, the adjustments to the periodic structure of existing timetable should also be considered guardedly.

3.2.2 Travel Time

An important factor for customer satisfaction is the total journey time which is a key aspect determining the attractiveness of a schedule.

The objective that aims at offering customers fast travel times should correspond to the dwell and connection requirements. However, rather than just satisfying these requirements, the objective is to satisfy them as well as possible. In other words, on the premise that there is sufficient dwell time for passengers to alight and board, the requirement of providing short travel time should be incorporated into the timetable as well as possible.

In addition, since the train speed is flexible in this paper, there are three decision variables influencing the total journey time for a passenger that are train speed, connection times and dwell times. Note that, due to the high speed in HSR, the deceleration and acceleration times at each braking and starting can not be omitted for providing a comfortable service to passengers. Consequently, the largest optimization potential here comes from the decrease of the times of stops and the waiting time at each stop.

3.2.3 Robustness

An other important factor for customer satisfaction is the robustness of a timetable. In a heavily utilization system that trains just meets the safety requirements in the timetable and may follow one another at exactly the minimum headway time, a small delay of one train may then be easily knocked on to other trains, and also propagated through the entire network. Due to large scope time period and many scheduled trains involved in the ATP problem, it is one of challenging problems for achieving a reliable railway system. Therefore, a important objective for the ATP problem is to insert the additional trains with the consideration that reconstructs a robust new timetable that contains some buffer time above the minimum headway time to absorb small disturbances.

While planning a timetable, certain buffer times are generally already added to the minimum headway time. Still, Leon (2003) defined that when track capacity is available, the timetable robustness is increased by “pulling apart” the trains as far as possible, since a delayed train is then less likely to interfere with the other trains on the track.

3.2.4 Makespan

Another possible objective for the ATP problem in China’s HSR is to minimize the makespan of the new timetable. As mentioned previously, in China’s HSR, the maintenance standard is at a very high level and needs to have a sufficient “wide” time interval within the cycle time, typically in the night from 0:00 to 4:00, during which no train is running. This disconnects the continuous diagram time and causes low track density during the early hours and late hours in the timetable. It should be noted that numerous schedules may have the same level of disruption and speed but completely different makespans. For example, if only constrained by the objective of travel time or deviation, the locations for additional trains that without time windows is always preferred in the early or late hours which of low traffic density in order to get a better result on disruptions and travel time. Consequently, the minimisation of disruptions or travel time are not necessarily sufficient as a single objective criterion for this problem.

3.2.5 Costs

An obvious fifth objective is to minimize the costs associated with the timetable. The major cost components of a railway system are formed by the infrastructure, the train-sets and the train crews. Given the assumption in the previous section, the infrastructure is fixed a prior, then the only few choices left that influence these operating costs of a railway system are train-sets and train crews schedules. In addition, train-set has direct implications for the passenger service and involves large amounts of money. The availability of appropriate quantitative models for supporting long term rolling stock management is highly important in practice. Although train-set planning and crews

scheduling usually occur after the timetable determined. Within limited freedom, we can still pursue the objective that insert additional trains with a minimum number of required train-sets.

3.3 Modeling Requirements

The following requirements of timetable in ATP problem should be taken into account:

3.3.1 Trip Time and Dwell Time Constraints

Trip time constraints are closely related to dwell time constraints, since together they model the journey time of a train from original to destination station.

We consider *variable trip times* in this paper to enlarge the solution space. The precise running times per train line enables to determine the percentile of running times within a certain level of probability. The positive difference between the scheduled running time and the technically minimal running time is called running time supplement. The typical time supplement is 3-7% on European railways depending upon speed of the trains. The higher percentage goes for high speed trains. In Netherlands, the running time supplements is about 7% and few minutes are added for recovery time at important locations in the train routes. In this paper, we apply 10% running time supplements in China's HSR. The time supplement also can be used to make up for delays during the extra train insertion.

In a timetable, the dwell times comprises scheduled dwell times at the stations calling at, and non-scheduled waiting times in practical operations. In the process of inserting trains, the dwell time at scheduled stops can be extended or compressed by certain extent respecting to the necessary time for passengers. Non-scheduled waiting times usually emerge from technical failures of track infrastructure for conflict resolution. In China's HSR, the minimum dwell time for each scheduled stop is set to 1-2 min, and taking the commercial reason into account, the maximum dwell time is bounded to 7 min.

Still, the trip times and dwell times are preferred to be as small as possible, that is, the train are preferred to run as the highest possible speed. This preference for small trip times and dwell times can be expressed by the objective function described later, which favors small travel time.

3.3.2 Safety Headways

The safety constraints ensure that the minimum headway time between trains is respected, and that the trains do not overtake each other on a track. Headways are the minimum time or distance by which the two consecutive running trains are separated in order to maintain safe operations. Generally, headways are controlled by signals which

show that if the trains are allowed to enter specified segments or it has to slow down or stop.

Hansen and Pachl (2008) and Pachl (2009) introduce a *blocking time model* to describe the headway in detail. With the blocking time stairway, it is possible to determine the minimum headway between two successive trains. Sufficient headway between trains is an essential safety requirement of timetabling. Two different ways to assign the headways to the station or to the section are adopted, as show in Figure 3.2 (Hansen and Pachl (2008)).

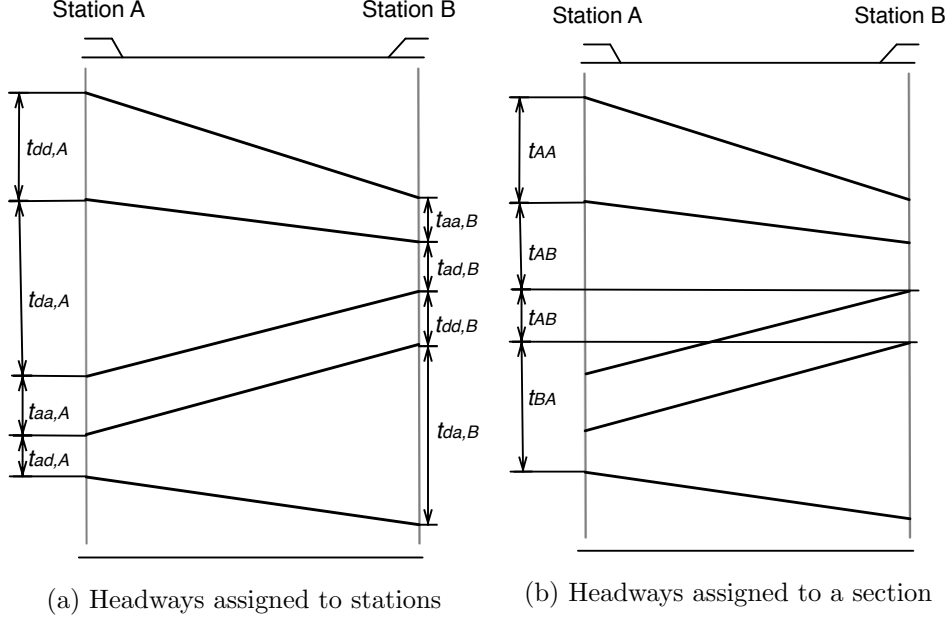


Figure 3.1: Principles of assigning headways

Assigning headways to stations is the more traditional principle and still common since it was introduced by Potthoff (1980), as shown in Figure 3.1a. There are four types of headways:

- (1) *Depart-Depart Headway* the headway between two successive trains that depart onto the same section, denoted as $t_{dd,A}$ and $t_{dd,B}$,
- (2) *Arrive-Arrive Headway* the headway between two successive trains that arrive from the same section, denoted as $t_{aa,A}$ and $t_{aa,B}$,
- (3) *Arrive-Depart Headway* the headway between arrival and departure of two successive trains that of opposite directions and toward the same section, denoted as $t_{ad,A}$ and $t_{ad,B}$,
- (4) *Depart-Arrive Headway* the headway between departure and arrival of two successive trains that of opposite directions and toward the same section, denoted as $t_{da,A}$ and $t_{da,B}$,

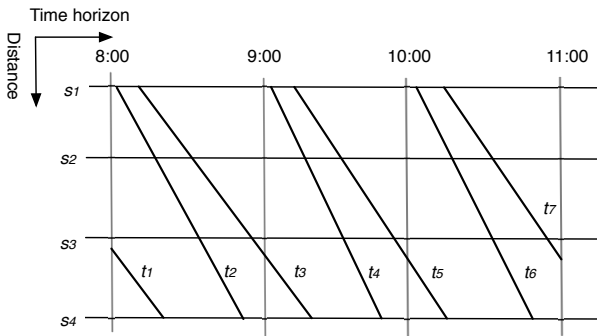
In this paper, we use this classical headway assignment as safety constraints.

3.3.3 Periodic Constraints

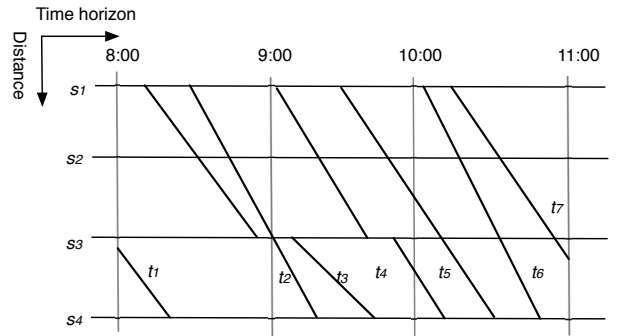
Since a significant factor for customer satisfaction is the cyclic behaviour of the train arrivals and departures in the cyclic timetable, keeping the periodic feature during the insertion is extremely important in the ATP problem.

Besides minimising the adjustments to existing trains which are scheduled periodically, the strategy of *cyclical rescheduling* or *periodic rescheduling* is essential to make up the deviations to the periodic structure. Periodicity rescheduling means that when a difference made to a cyclic train, the trains in other cyclic times which belong to the same train line with the disturbed trains should also be rescheduled to ensure the cyclic arrivals and departures in the entire timetable. In contrast to the conventional rescheduling strategy which is common in dispatching problem in previous researches, the cyclical rescheduling takes the periodic structure into consideration besides resolving conflicts. However, in practical applications we do not want to fixed the reschedules too much beforehand, then a bandwidth is introduced to the periodic constraint.

For example, in a cyclic timetable shown in Figure 3.2a, there exist two different cyclic train lines $\{t_1, t_3, t_5, t_7\}$ and $\{t_2, t_4, t_6\}$, respectively. Note that the non-cyclic trains are omitted for clarity. Due to insertion, train t_2 have to left shift by 30 min. If the conventional rescheduling strategy is adopted, the new timetable after conflict resolution is shown in Figure 3.2b. The periodic structure of the timetable is totally violated. If the cyclical rescheduling without a bandwidth (i.e. bandwidth=0) is applied, along with adjusting t_2 , t_4 and t_6 which belong to the same train line with t_2 but in different cyclic times are also rescheduled by 30 min to keep an exact periodic structure, as shown in Figure 3.2c. While if a bandwidth of 5 min is introduced, t_4 and t_6 can be shift by $[30 - 5, 30 + 5]$ min, then the new timetable is shown in Figure 3.2d. Compared with the solution in Figure 3.2c, a bandwidth can decrease the affected number of trains and the total adjustments, and simultaneity keep the periodic structure in an acceptable level.



(a) A cyclic timetable



(b) Conventional rescheduling

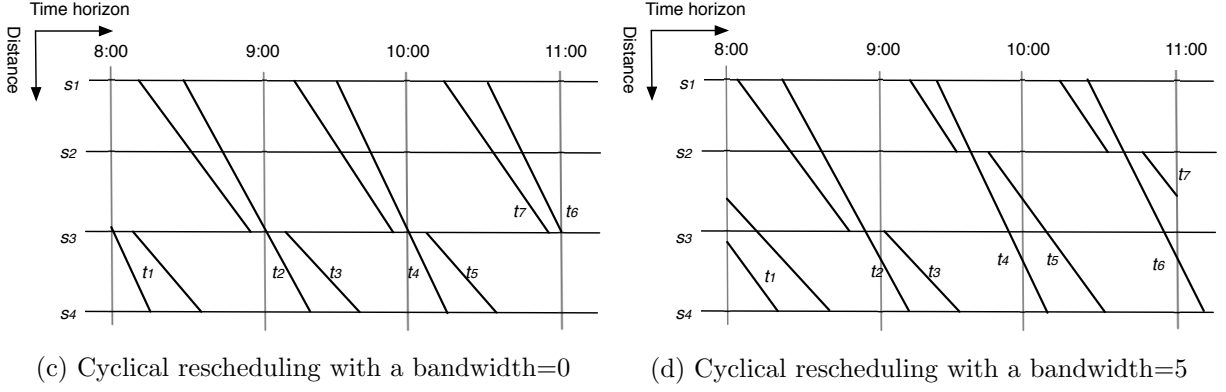


Figure 3.2: Various strategies of rescheduling

3.3.4 Connection Constraints

Connections are composed of transfer connections and combination constraints.

A transfer connection allows passengers from the predecessor train transfer to the successor train at a station. A connection time is used for the connections between these two trains. In China, due to the massive passengers, the minimum transfer time is usually set to 15 min.

A combination connection allow two trains from different directions combined into one train at a station. However, in China's HSR only the train-sets of 8 passenger cars and 16 passenger cars are used; combining and splitting are not considered in practical applications, which will be further described in Section 5.2.1. Therefore, only transfer connection constraints are modelled in this paper.

3.3.5 Frequency Constraints

Leon (2003) defines the synchronization constraints in trains from different train lines share part of their routes. When this synchronization of trains within a train line, the involved constraints are also known as *frequency constraints*.

Here, we use the definition of synchronization constraints in Leon (2003) to explain the frequency constraints in this paper. Frequency constraints are applied to spread the multiple trains of a single train line evenly across the considered time horizon. As an example, if two trains of the same services need to be inserted in one hour, then synchronizing their departures provides a service with frequency two on the common part of their route. However, we do not want to fix the timetable too much beforehand. Therefore the model requires the departure times to be 30 minutes apart, with a bandwidth of two minutes. Note that the 30 minutes that the trains should lie apart is obtained by dividing the consider time horizon of 60 min by the frequency of two.

3.3.6 Time Window Constraints

Because of commercial reasons, the departures and arrivals of passenger trains are often be constrained within predefined time windows, especially at the origin and destination station. Time windows are given to both existing services and to new services. All of the services should satisfy preferred time windows if they have been defined. It should be noted that for various types of train, the freedom of selecting departure time is also different, especially for international trains and cross-line trains.

3.3.7 Station Capacity

By viewing a station as a node, the details of tracks within station are lost and stations are assumed to be black boxes.

In theory the station capacity may be modelled explicitly or implicitly. An explicit approach distributes a exactly track within station for each train. This causes additional routing alternatives (routing flexibility) with respect to number of tracks and no-conflict route. Leon (2003) introduces the problem of routing trains through station tracks usually is assumed to be solved in a later phase of the timetable planning process. Lusby et al. (2011a), Lusby et al. (2011b) and Nie and Hansen (2005) describe the in-station routing of trains in more detail.

An implicit approach represents the station as a capacitated intermediate storage area, and a feasible timetable resulting from the model should respect to the capacity of station (typically denoted by the number of station tracks). In Section 4.1, we relax this black box station assumption to a certain extent and present an adding paths model that allows for incorporating station capacity into the inserting and dispatching, which results in a grey box.

Chapter 4

Mathematical Model for Adding Train Paths Problem

This chapter describes a model for the ATP problem. Section 4.1 describes the general mixed integer program (MIP) model for the ATP problem. In this work, the problem is characterized based on an event-activity graph. Section 4.2 formulates several additional real-world constraints which deal with the acceleration and deceleration time, priority for overtaking, station capacity, frequency services, allowed adjustment and allowed deviations to periodic structure to initial timetable. Next, Section 4.3 formulates the considered objectives (i) minimizing travel time of additional trains, (ii) minimizing total adjustments to initial trains, (iii) minimizing the makespan and (ix) maximizing the robustness of the new timetable, in order to get a new timetable that with quality of the performance to the additional trains, low deviations to the initial services and high quality of the entire trains, respectively. Finally, a helpful adding train paths tool is developed and Section 4.4 uses this tool to test the proposed techniques based on the Shanghai-Hangzhou HSR line in China. The experiments are divided to two part, that the first part investigates the influence of using different objective functions and frequency on the values of identified performance measures, while the second part analyses if and how various tolerance of disruptions to initial timetable influences the insertion effect.

4.1 A General Adding Train Paths Model

This section formulates a general mixed integer program (MIP) model for the ATP problem. This ATP model is described based on the event-activity graph.

4.1.1 Railway Network Input

The sets below contain the basic information for the railway timetable.

$\mathcal{G} = (\mathcal{S}, \mathcal{B})$	The railway network graph, consisting of stations \mathcal{S} , and sections \mathcal{B} .
\mathcal{S}	Each station $s \in \mathcal{S}$ is denoted as a station for train's arrival, departure, stop, overtaking and so on.
\mathcal{B}	Each section $b \in \mathcal{B}$ is denoted as a collection tracks between stations, with no intermediate station in between.
\mathcal{T}	The set of all trains.

4.1.2 An Event-activity Graph

In order to profit from specific topological structures of each individual railway network, an event-activity graph $G = (V, E)$ is adopted in this paper, which is firstly used by Schöbel (2001a) in railway timetabling.

An event-activity network $G = (V, E)$ is a directed graph whose nodes V are called *events* and whose directed edges E are called *activities*. Event-activity networks are a widely used mathematical model for cyclic or acyclic scheduling of events with time constraints. In the acyclic case which we consider here, an activity which connects two events models a precedence constraint between those events. Each activity has assigned a lower bound on its duration, so the scheduled time of the end event of an activity has to be larger than or equal to the scheduled time of the start event plus the lower bound. In contrast to the cyclic case, in cyclic event-activity networks (used for example for cyclic timetabling), each activity has assigned a lower and an upper bound, modeling time window constraints.

The set V of events consists of all *arrival events* and *departure events*, i.e. $V = V_{arr} \cup V_{dep}$,

$$\begin{aligned} V_{arr} &= \{(t, s, arrival) : \text{train } t \in \mathcal{T} \text{ arrives at station } s \in \mathcal{S}\}, \\ V_{dep} &= \{(t, s, departure) : \text{train } t \in \mathcal{T} \text{ departs from station } s \in \mathcal{S}\}, \end{aligned}$$

The events of set V are linked by directed edge set E , which are called *activities* and consists:

- *Trip activities* $E_{trip} \subset V_{dep} \times V_{arr}$ model the travelling time of a train between two consecutive stations, so a trip activity connects a departure event of a train with its next arrival event at the subsequent station. The lower bound $trip_e^{min} > 0$ and upper bound $trip_e^{max} > 0$ of a trip activity $e \in E_{trip}$ represents the minimal and maximum driving time respectively between both stations. When $trip_e^{min} = trip_e^{max}$, the trip time of train between consecutive stations is assumed to be fixed.
- *Dwell activities* $E_{dwell} \subset V_{arr} \times V_{dep}$ model the the stopping time of a train within a station, such as boarding and de-boarding of passengers or for crew change. A dwell activity connects the arrival of a train at a station with its departure from

the same station. The lower bound $dwel\ell_e^{min} > 0$ of a dwell activity $e \in E_{dwell}$ describes the minimal time which is needed to let passengers get on or off and also takes into account the time for crew change or other actions. In contrast, the upper bound $dwel\ell_e^{max} > 0$ of a dwell activity represents the maximum dwell time when the rapidness of travel for passenger trains is taken into consideration. When $dwel\ell_e^{min} = dwel\ell_e^{max} = 0$, train passes through the station without stop. Each activity in E_{trip} and E_{dwell} corresponds to an action of one train. We summarize them in set $E_{train} := E_{trip} \cup E_{dwell}$. Consider a train t on the journey from its origin station to its destination station. From Figure 4.1, it is clear that the total journey of a train t defines a path P_t in G , starting and ending with a trip arc, and in between consisting alternatively of dwell arcs and trip arcs.

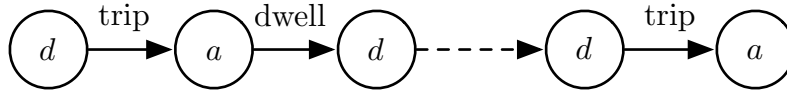


Figure 4.1: Train path P_t and its corresponding activities E_{train}

- *Changing activities* $E_{change} \subset V_{arr} \times V_{dep}$ model a transfer connection from one station to another. It allows passengers to transfer from train t to train t' within the same station, so a changing activity connects an arrival event of train t with a departure event of train t' at the same station. The lower bound $change_e > 0$ refers to the minimum time the passengers need when they transfer between these two trains. It is one of the tasks of adding paths problems to decide for each changing activity if the corresponding connection should be maintained or not. If a connection is maintained, the lower bound $change_e$ of the corresponding changing activity $e \in E_{change}$ has to be respected, otherwise it can be ignored.
- *Headway activities* $E_{headway} \subset V_{dep} \times V_{dep} \cup V_{arr} \times V_{arr}$ model the security headway between two consecutive departures and arrivals at the same station. $E_{headway}$ is a set of headway arcs of the form $\langle (v_i, v_j), (u_i, u_j) \rangle$ and it models the limited capacity of the track system. Consider two trains t_1, t_2 travelling on a same section $b = (s_1, s_2)$, see Figure 4.2. Event v_1 represents the departure of t_1 from the station of s_1 and u_1 is its arrival at the station of s_2 . Similarly, event v_2 represents the departure of t_2 from station s_1 and u_2 is its arrival at the final station of s_2 . As analysed in section 4.1.1, a minimum time distance between two consecutive arrivals h_a and departures h_d has to be respected. Besides, due to overtaking along a section is not allowed, the pair of arcs (v_1, v_2) and (u_1, u_2) , say $\langle (v_1, v_2), (u_1, u_2) \rangle$, or (v_1, v_2) and (u_1, u_2) , say $\langle (v_2, v_1), (u_2, u_1) \rangle$, should be selected simultaneously. (v_1, v_2) and (v_2, v_1) are alternative arcs pair, and exactly one headway activity from each alternative pair has to be respected. If (v_1, v_2) is chosen, then train t_1

departures before t_2 from station s_1 and (u_1, u_2) has to be respected in order to prevent collision on section b . On the contrary, if (v_2, v_1) is chosen, then train t_2 departures before t_1 from station s_1 and (u_2, u_1) must be selected. The goal of adding paths problem hence is to choose exactly one activity of each such pair and to respect the resulting constraint, fixing the order of the two trains to occupy a same section.

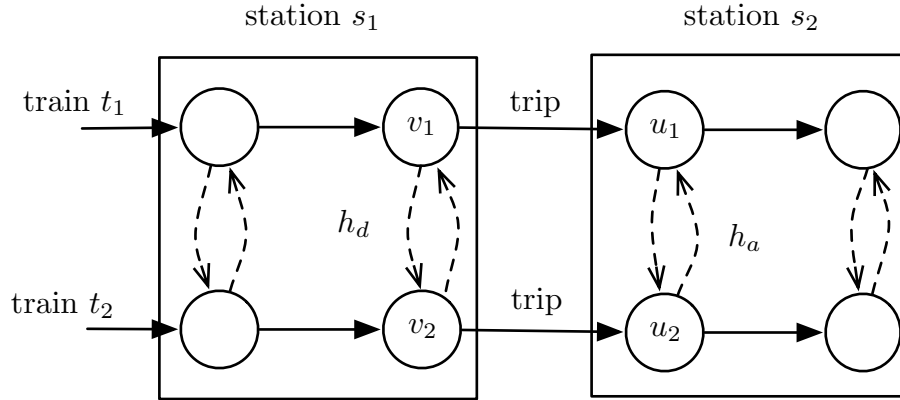


Figure 4.2: Headway arcs between two consecutive trains on a same section

To illustrate an event-activity network, we use the following example (which is depicted in Figure 4.3). Figure 4.3a sketches a small railway network with 3 stations and 3 trains. train 1 and train 2 drive from station a to station b , while train 3 drives from station a to station c . Within station a , passengers might transfer between trains 1 and 3, and trains 2 and 3. Within station b , passengers transfer between trains 1 and 2.

Figure 4.3b shows the corresponding event-activity graph G . For each station, a rectangle is drawn around the nodes corresponding to that station. The solid lines and dotted lines in the constraint graph represent the trip arcs, and dwell arcs respectively.

All safety arcs are dashed. When the trip times of trains between consecutive stations were assumed to be fixed, it suffices to state the safety constraints for the departures of trains only. In contrast, when variable trip times is incorporated, the minimum intervals between departures and arrivals have to be restricted.

The connection arcs have been drawn in bold. Since the three trains head for two different destinations after leaving station a , connections are defined between these trains of different destinations. So passengers can travel from any of train 1 and 2 to train 3 in station a with a good connection. Similarly, passengers of train 1 and train 2 can be exchanged in station b .

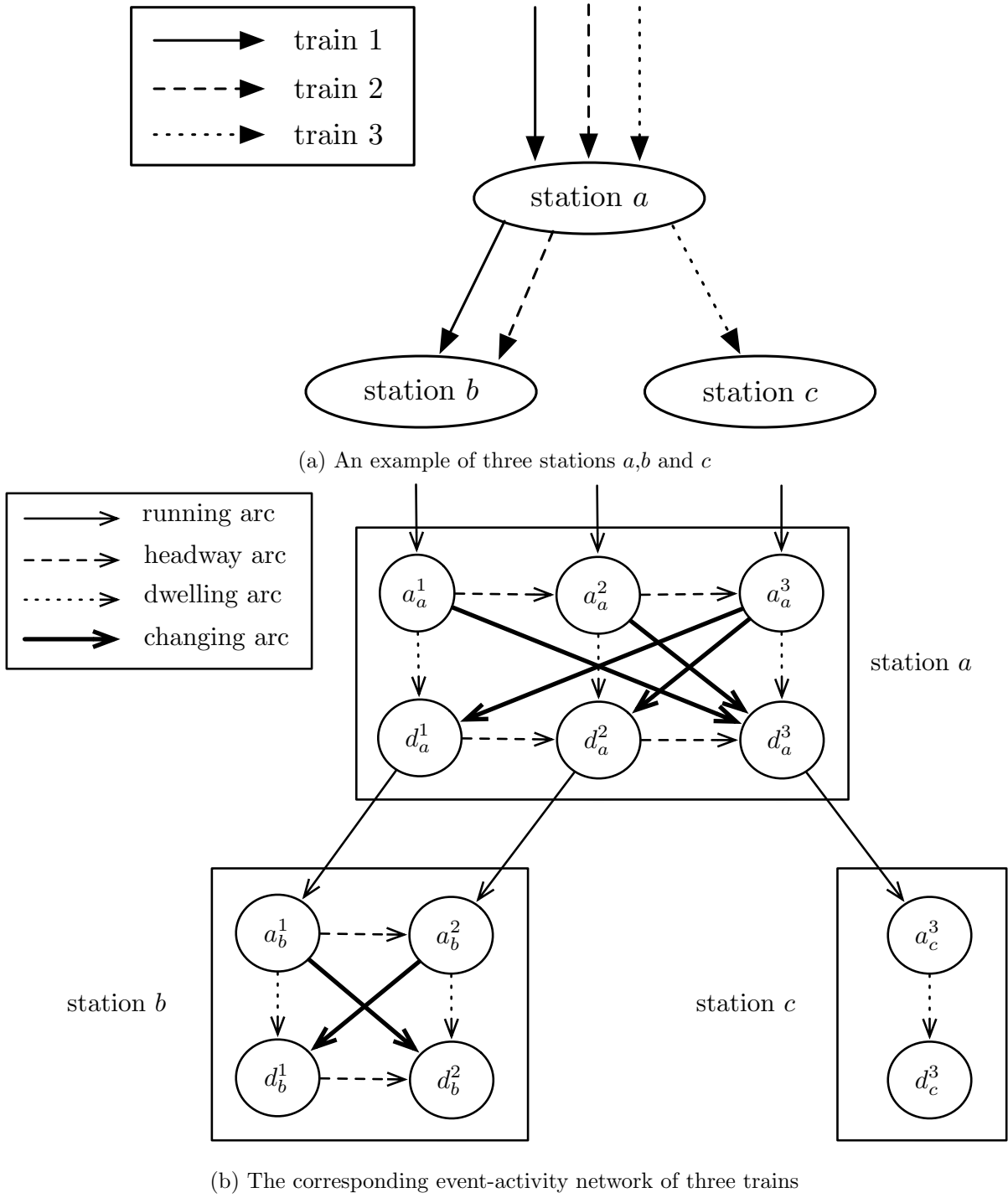


Figure 4.3: A small railway network with three stations and three trains, and the corresponding event-activity graph

4.1.3 Model and Integer Programming Formulation

Problem input: trains and timetables

The set of trains considered is given by $\mathcal{T} = \mathcal{T}^{ini} \cup \mathcal{T}^{add}$, where \mathcal{T}^{ini} denotes the sets of initial trains (generally cyclic trains) that have a prescribed timetable and \mathcal{T}^{add} denotes

the set of additional trains (generally non-cyclic trains) that need to be inserted to the initial timetable.

For each train $i \in \mathcal{T}^{ini}$, a timetable is specified, consisting of:

- an ordered sequence of stations $\mathcal{S}^i := \{f_i, \dots, l_i\} \subseteq \mathcal{S}$ that the train i visits, where f_i is the first (origin) station and l_i is the last (destination) station;
- the departure time from f_i , the arrival time at l_i , and the arrival and departure time for the intermediate stations in $\mathcal{S}^i \setminus \{f_i, l_i\}$;
- the exact track k for the allocation of train i on each section and station;
- the rolling-stock composition circulation.

For each train $i \in \mathcal{T}^{add}$, it is predefined as following:

- an ordered sequence of stations $\mathcal{S}^i := \{f_i, \dots, l_i\} \subseteq \mathcal{S}$ that the train i visits, where f_i is the first (origin) station and l_i is the last (destination) station;
- the desired departure time window from station f_i , the minimum dwell time at each station in $\mathcal{S}^i \setminus \{f_i, l_i\}$ and the trip time window at each section $b = (h, j)$, with $h, j \in \mathcal{S}^i$;
- the desired transfer connections involved with additional trains;
- the number of available rolling stock compositions for additional trains.

Notations

Table 4.1 and Table 4.2 first list general subscripts and input parameters used in the proposed model. Table 4.3 describes the decision variables in the proposed optimization model. The unit of all time-related parameters and variables is one minute. In this study, we focus on a train timetabling problem on a double-track rail line which consists of a series of uni-directional track segments. Two types of trains with different priorities (i.e. high-speed trains and medium-speed trains) traverse on the rail corridor.

Table 4.1: General subscripts and parameters

Symbol	Description
V^{add}	= set of additional events
V^{ini}	= set of initial events
V_{dep}	= set of departure events
V_{arr}	= set of arrival events
i, j	= event index
$\sigma(i)$	= the successor event of i
e	= activity index, $e = (i, j)$
$s(i)$	= the station at which event i takes place

(Continued on next page)

Table 4.1 – (continued from previous page)

Symbol	Description
$b(i, j)$	= the section on which activity $e = (i, j)$ takes place
$t(i)$	= the train of event i
$chang_e$	= the minimum time for transfer connection
ϵ^a	= required acceleration time *
ϵ^d	= required deceleration time *
h	= The minimum headway time between two consecutive events *
p_{ij}	= 1 if train $t(i)$ has higher priority than train $t(j)$ = 0 otherwise
U_s	= the number of tracks in station s
M	= A sufficiently large positive integer

* We use general acceleration time ϵ^a , deceleration time ϵ^d and headways h here. If desired, one can also specify these parameters that depend on the type of the involved trains and stations. For example, the headway between two faster trains may usually be larger than the headway between two slower trains.

Table 4.2: Input parameters for inserting train paths

Symbol	Description
π_i	= the time instant at which event $i \in V$ takes place in the initial timetable
pls_i	= 1 if train $t(i)$ stops at the station $s(i)$ in the initial timetable = 0 if train $t(i)$ bypasses the station $s(i)$ in the initial timetable
tw_i^{min}	= the lower bound of the time window at which event i takes place
tw_i^{max}	= the upper bound of the time window at which event i takes place
$trip_e^{min}$	= the minimum trip time of event e
$trip_e^{max}$	= the maximum trip time of event e
$dwell_e^{min}$	= the minimum dwell time of event e
$dwell_e^{max}$	= the maximum dwell time of event e
Δ_i	= maximum allowable adjustment of event i
T	= cyclic time, 1 hour in this paper
N	= number of additional trains
θ	= maximum allowable deviation to periodic structure
β	= bandwidth of frequency
T_{hor}	= considered time horizon

Table 4.3: Decision variables.

Symbol	Description
x_i	= the time instant at which event $i \in V$ takes place in the new timetable
λ_{ij}	= 1 if event j takes place after, or at the same time as event i , = 0 if event j takes place before event i
ρ_i	= 1 if train $t(i)$ stops the station $s(i)$ = 0 if train $t(i)$ bypasses the station $s(i)$
y_e	= 0 if the connection e is kept = 1 otherwise

Constraints

In the following, we use the concept of event-activity networks to give a mathematical formulation of the add paths problem. Constraints used in the double-track adding paths model are presented as the following,

Reasonable time window:

$$x_i \geq tw_i^{min} \quad \forall i \in V \quad (4.1)$$

$$x_i \leq tw_i^{max} \quad \forall i \in V \quad (4.2)$$

Variable trip time on section:

$$x_j - x_i \geq trip_e^{min} \quad \forall e = (i, j) \in E_{trip} \quad (4.3)$$

$$x_j - x_i \leq trip_e^{max} \quad \forall e = (i, j) \in E_{trip} \quad (4.4)$$

Dwell time at station:

$$x_j - x_i \geq \rho_i \cdot dwell_e^{min} \quad \forall e = (i, j) \in E_{dwell} \quad (4.5)$$

$$x_j - x_i \leq \rho_i \cdot dwell_e^{max} \quad \forall e = (i, j) \in E_{dwell} \quad (4.6)$$

$$\rho_i = 1 \quad \forall i \in V_{arr} : pls_i = 1 \quad (4.7)$$

Transfer connection time:

$$x_j - x_i \geq change_e \quad \forall e = (i, j) \in E_{change} \quad (4.8)$$

$$x_j - x_i \geq change_e - My_e \quad \forall e = (i, j) \in E_{change} \quad (4.9)$$

Minimum headway:

$$x_j - x_i \geq h_e \cdot \lambda_{ij} - M \cdot (1 - \lambda_{ij}) \quad \forall (i, j) \in E_{headway} \quad (4.10)$$

$$x_i - x_j \geq h_e \cdot (1 - \lambda_{ij}) - M \cdot \lambda_{ij} \quad \forall (i, j) \in E_{headway} \quad (4.11)$$

$$\lambda_{ij} = \lambda_{\sigma(i)\sigma(j)} \quad \forall b(i, \sigma(i)) = b(j, \sigma(j)) \quad (4.12)$$

Operator preferences:

$$x_i \geq 0 \quad \forall i \in V \quad (4.13)$$

$$\rho_i, \lambda_{ij} \in \{0, 1\} \quad \forall i \in V \quad (4.14)$$

- Constraints (4.1-4.2) represent the reasonable departure time window for trains. For some trains, the freedom of selecting departure times from original station (arrival times at destination) is limited. This especially applies to international trains and interline trains. Time windows of departure (arrival) times are usually chosen on the board stations. When $tw_i^{min} = tw_i^{max}$, it ensures that the departure (arrival) time is fixed. When $tw_i^{min} = \pi_i, \forall i \in V^{ini}$, it ensure that no event takes place earlier than scheduled in the original timetable.

- Constraints (4.3-4.4) relate the actual trip time on section. Taking speed variation dynamics into consideration, the trip time in section is flexible between the minimal $trip_e^{min}$ and the maximal $trip_e^{max}$. When speed is fixed on sections, then $trip_e^{min} = trip_e^{max}$.
- As shown in constraints (4.5-4.7), train must stop at all stations at which it calls (i.e. $pls_e = 1$, else $pls_e = 0$). More precisely, extension of a scheduled stop or additional stops is permitted for operational requirements. Due to commercial and operating reasons, stopping time must be bounded. The actual dwell time should be no less than the planned minimum $dwell_e^{min}$ and no more than the maximum $dwell_e^{max}$ dwell time.
- Constraints (4.8) forces connecting trains have to wait, while in heavily delayed situations. Since a train may not be able to wait for a train delayed more than a certain time, the waiting condition should ideally be dynamic and dependent on how delayed the connecting train is. (4.8) are hard connection restrictions. In contrast, constraints (4.9) represent a possible relevant extension about varying of connections, and it is only relevant if objective function

$$\min \sum_{e \in E_{change}} y_e$$

is used. If a connection $e \in E_{change}$ is kept, then $y_e = 0$ and the corresponding constrain ensures that the lower bound on the duration of this connecting activity is respected. However, when the connection is missed, (4.9) impose no additional constraints.

- The headway constraints (4.10-4.12) guarantee the safety headways between trains. Considering each type of potential conflict between each pair of trains, we impose a specific separation time h between departure and/or arrivals of the two trains. As headway constraints are alternative arcs pair (either event i takes place before event j or event j takes place before event i), constraints (4.10-4.11) make sure that exactly one headway arc are selected from each pair, and they together model the constraints:

$$x_j - x_i \geq h_e \vee x_i - x_j \geq h_e$$

Constraints (4.12) make sure that overtaking on the track section must be avoid between any two trains, i.e. the order of departures and arrivals of any two trains can not be changed on a track section.

4.2 Extensions of the Adding Train Paths Problem Model

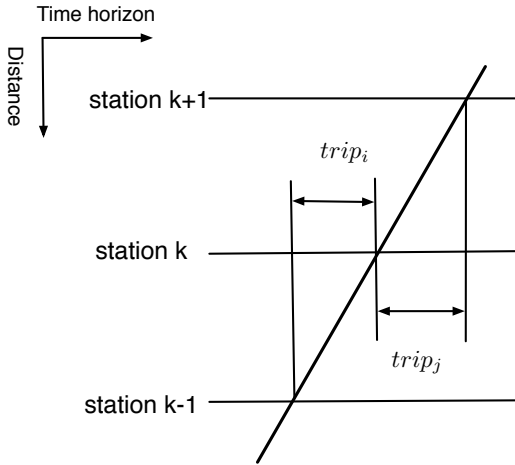
The goal of this section is to present a model for the ATP problem with several real-world constraints in which the decision dealing with the acceleration and deceleration time, priority for overtaking, station capacity, frequency constraints and tolerance of disruption for initial timetable which consists of allowed adjustment and allowed deviations to the periodic structure are integrated.

4.2.1 Acceleration and deceleration time

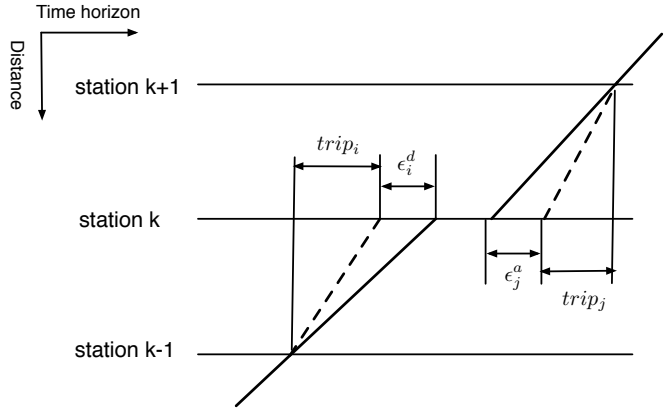
In addition to the flexible velocity, due to the safety and passenger comfort requirements, the high-speed trains usually take at least several minutes to fully stop or reach a cruise speed even with highly efficient acceleration and deceleration performance (Zhou and Zhong (2005)). In this situation, when train stops the corresponding actual trip time has to exactly take into account the required acceleration time ϵ^a and deceleration time ϵ^d .

Figure 4.4 shows that if train t stops at station k , the required times for acceleration and deceleration at station k have to be set to ϵ_i^a and ϵ_j^d respectively, where $i = (t, k, \text{departure}) \in V_{dep}$ and $j = (t, k, \text{arrival}) \in V_{arr}$. From the resource-constrained project scheduling point of view, time is the discrete resource in this problem, and there are four execution modes for train t travelling in section $b = (k, k + 1)$, namely,

- (1) bypasses both stations k and $k + 1$,
- (2) stops at station k and bypasses at station $k + 1$,
- (3) bypasses at station k and stops at station $k + 1$,
- (4) stops at both stations k and $k + 1$.



(a) Bypasses station k



(b) Stops at station k

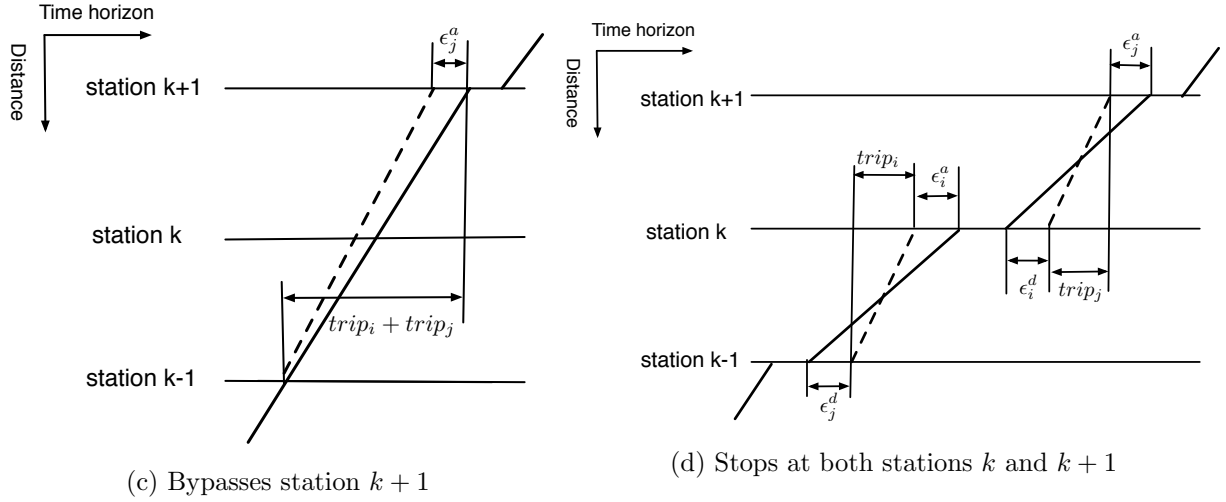


Figure 4.4: Illustration for acceleration time and deceleration time

Then constraints(4.3-4.4) is rewrote by constraints (4.15-4.16) which imply that the acceleration and deceleration times at station k and $k + 1$ are added to the trip time in section $b = (k, k + 1)$.

Acceleration and deceleration time:

$$x_j - x_i \geq trip_e^{min} + \rho_i * \epsilon_i^a + \rho_j * \epsilon_j^d \quad \forall e = (i, j) \in E_{trip} \quad (4.15)$$

$$x_j - x_i \leq trip_e^{max} + \rho_i * \epsilon_i^a + \rho_j * \epsilon_j^d \quad \forall e = (i, j) \in E_{trip} \quad (4.16)$$

4.2.2 Priority for Overtaking

Since the capacity of track system is limited, usually many conflicts arise in case of insertions of additional trains or shifts of existing trains. If two trains are competing for the same section, the decision has to be made that which train is allowed to go first and which train has to wait. The above described a prior assignment of trains to sections and also applied to station side tracks. Usually, it should be decided beforehand which trains should move to a side track to be overtaken. The dispatcher attempts to keep the negative impact of existing train services as low as possible, and tries to prevent knock-on conflicts arising, of at least to minimise them.

In order to reach these goals, the dispatcher uses the following dispatching rules which have been proven in practice (Hansen and Pachl (2008)):

- emergency trains get highest priority,
- premium trains are prioritised to other trains,
- fast trains get preference over slow trains,
- dedicated lines offer certain trains priority over other trains (e.g. freight trains on freight lines)

For sake of simplicity, velocity is used to distinguish the priority of trains in this paper. When a overtaking conflict arises, it is cheaper to stop the slow train at station and wait for a period time. In addition, if the train being stopped is the fast train rather than the slow train, then the fast train (and all trains behind) will have to wait for a very long time, or stop again and again in the following stations to avoid the possible overtakings with the slow train, which is not desirable. So the slow train with low priority is permitted to be overtaken by fast train with high priority, and a fast train is forbidden to be overtaken by a slow train or another fast train with the same priority.

More precisely, let p_{ij} indicate the relative priority between trains $t(i)$ and $t(j)$ as follows:

$$p_{ij} = \begin{cases} 0 & \text{if train } t(i) \text{ has lower, or same priority than train } t(j) \\ 1 & \text{if train } t(i) \text{ has higher priority than train } t(j) \end{cases} \quad (4.17)$$

Consider station s that has more than one side track for overtaking action. The overtaking constraints are defined by specifying two trains with the arrival events $i, j \in A$ and the corresponding departure events $\sigma(i), \sigma(j) \in D$ at the same station s respectively, shown as in Figure 4.5.

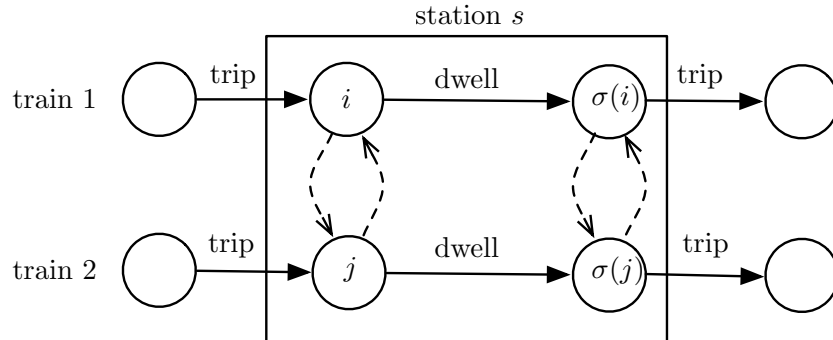


Figure 4.5: Train priority at station

As described in Section (4.1.3), we say that train $t(i)$ arrives before train $t(j)$ (i.e. $i \prec j$) if $\lambda_{ij} = 1$. Analogously, we will use the notation $i \succ j$, $\sigma(i) \prec \sigma(j)$, $\sigma(i) \succ \sigma(j)$.

Theorem 4.2.1. *Consider two trains as described above. Then the following constraints ensure that the overtaking rule is respected:*

$$p_{ij}(1 - \lambda_{ij}) + (1 - p_{ij})\lambda_{ij} \geq p_{ij}(1 - \lambda_{\sigma(i)\sigma(j)}) + (1 - p_{ij})\lambda_{\sigma(i)\sigma(j)}$$

Proof. According to the dispatching rule, all of the potential overtaking are illustrated as follows:

- When train $t(i)$ has higher priority than train $t(j)$,

- if $i \prec j$ (i.e. $\lambda_{ij} = 1$), since the faster train $t(i)$ can not be overtaken by the slower one, then $\lambda_{\sigma(i)\sigma(j)} = \lambda_{ij} = 1$.
- if $i \succ j$ (i.e. $\lambda_{ij} = 0$), since the slower train $t(j)$ could be possible overtaken by $t(i)$, then $\lambda_{\sigma(i)\sigma(j)}$ has two values $\{0, 1\}$. When $\lambda_{\sigma(i)\sigma(j)} = 1$, $\sigma(i) \prec \sigma(j)$ and train $t(j)$ is overtaken by train $t(i)$.

then, $\lambda_{ij} \leq \lambda_{\sigma(i)\sigma(j)}$ for all $p_{ij} = 1$.

- When train $t(i)$ has lower priority than train $t(j)$,
 - if $i \prec j$ (i.e. $\lambda_{ij} = 1$), since the slower train $t(i)$ could be possible overtaken by $t(j)$, then $\lambda_{\sigma(i)\sigma(j)}$ has two values $\{0, 1\}$. When $\lambda_{\sigma(i)\sigma(j)} = 0$, $\sigma(i) \succ \sigma(j)$ and train $t(j)$ is overtaken by train $t(i)$.
 - if $i \succ j$ (i.e. $\lambda_{ij} = 0$), since the faster train $t(j)$ can not be overtaken by the slower one, then $\lambda_{\sigma(i)\sigma(j)} = \lambda_{ij} = 0$.

then, $\lambda_{ij} \geq \lambda_{\sigma(i)\sigma(j)}$ for all $p_{ij} = 0$

□

Then the constraints to decide which train should wait and which train should go first with priority decisions can be represented by,

Priority for overtaking:

$$p_{ij}(1 - \lambda_{ij}) + (1 - p_{ij})\lambda_{ij} \geq p_{ij}(1 - \lambda_{\sigma(i)\sigma(j)}) + (1 - p_{ij})\lambda_{\sigma(i)\sigma(j)} \quad \forall i, j \in V \quad (4.18)$$

4.2.3 Station Capacity

The ATP problem so far considered stations to be black boxes. It returns train arrival and departure time, and leaves the construction of feasible platform assignments and routing through the stations to be carried out in a later phase. This section described how to model the capacity of a station to some extent. The results are more generally applicable to model the capacity of a general node in the railway network.

Let U_s be the capacity of a station $s \in \mathcal{S}$, which expresses as the maximum number of the trains that can be in the station s at the same time instant in railway network $\mathcal{G} = (\mathcal{S}, \mathcal{B})$. Usually, U_s represents the number of tracks in a station, or more generally, the number of trains can be handled in a station at the same time. This section describe how to restrict a upper bound U_s on the number of trains that is concurrently existing at the corresponding station s . The basic idea is the following. For every train t , we count the train that are present at station s at the same time as t , and limit that number by the capacity U_s (Leon (2003)). The analysis is based on the interpretation of the variables λ_{ij} which is described in Section 4.1.3.

Suppose that a station graph with arrival events set A and departure events set D contains all arcs:

$$\begin{aligned} (i, j) \quad & \text{with } i, j \in A, \text{ and} \\ (i, j) \quad & \text{with } i \in D, j \in A \end{aligned}$$

The sequence between two arrival events $i, j \in A$ can be obtained from the headway constrain (4.10-4.11), and the sequence of every departure event $i \in D$ and every arrival event $j \in A$ can be restricted by:

$$M\lambda_{ij} + x_i - x_j \geq 0 \quad \forall i \in D, j \in A \quad (4.19)$$

$$M(1 - \lambda_{ij}) + x_j - x_i \geq 0 \quad \forall i \in D, j \in A \quad (4.20)$$

If event i takes place before event j , i.e. train $t(i)$ departs before the arrival of train $t(j)$, then $\lambda_{ij} = 1$. Otherwise, $\lambda_{ij} = 0$.

Then the following constraints ensure that a station capacity of U_s is respected:

Station capacity:

$$1 + \sum_{i \in A} \lambda_{ij} - \sum_{i \in D} \lambda_{ij} \leq U_{s(j)} \quad \forall j \in A \quad (4.21)$$

Consider a arrival event j and the corresponding train $t(j)$ and station $s(j)$. The first sum term in (4.21) represents train t itself. The second sum term in (4.21) counts the number of trains that arrive at station $s(j)$ before, or at the same time as train $t(j)$. Similarly, the third sum term counts the number of trains that leave $s(j)$ before train $t(j)$ arrives.

Thus, adding the trains arriving before the arrival of $t(j)$ and train $t(j)$ itself, and abstracting the trains that leaved before $t(j)$ arrives, the total number of trains present at station $s(j)$ concurrently with $t(j)$ can be obtained. For any train $t(j)$, the total number of trains in station $s(j)$ concurrently is limited by $U_{s(j)}$ (Leon (2003)).

4.2.4 Tolerance of Disruptions for Initial Timetable

During the procedure of insertion, initial timetable can not avoid to be adjusted in many situations. What an acceptable level of disruption is however is fairly subjective. For the adding paths problem described in this paper, the tolerance of disruption for initial cyclic timetable can be constrained by:

- (1) Allowed adjustments, and
- (2) Allowed deviations to the periodic structure.

Allowed adjustment

The level of acceptable adjustment widely differs according to the train service type. It should also be mentioned that in some circumstances some train services must be strictly fixed and can suffer no disruption, such as cross-line and some high-speed trains. Similarly, the tolerance of other passenger trains in China's HSR to delays and alterations is quite limited. In this paper, Δ_i is introduced as the maximum adjustment of event $i \in V^{ini}$ to constraint disruptions for initial trains. Clearly, $\Delta_i \geq 0$, and if the corresponding schedule is fixed then $\Delta_i = 0$ holds.

Allowed adjustment of initial schedules:

$$x_i - \pi_i \leq \Delta_i \quad \forall i \in V^{ini} \quad (4.22)$$

$$\pi_i - x_i \leq \Delta_i \quad \forall i \in V^{ini} \quad (4.23)$$

Constraints (4.22-4.23) equal to $|x_i - \pi_i| \leq \Delta_i$ and imply that only a certain amount of left or right shift are allowed for initial trains.

Allowed deviations to the periodic structure

Besides, the additional trains are inserted while taking the structure of the planned cyclic timetable into account. In a cyclic timetable, train connections are operated regularly with respect to a cycle time. So, a train for a certain destination leaves a certain station at exactly the same time every cycle time T . During the process of adding and adjusting the schedules of trains, one usually runs into problems that the periodicity of initial cyclic timetable might be ruined. In order to fully take the advantage of cyclic timetable, the periodic pattern of initial trains is desired to be guaranteed. However, sometimes we do not want to fix the initial timetable too much beforehand. Therefore the model requires the departure times to be T minutes apart, with a bandwidth of θ minutes.

Considering the periodic relation between the operations of train t_1 in the first time cycle, and the other trains t_2, \dots, t_k in the k -th cycle time, this relation is defined by the following constraints:

$$x_k - x_1 \in [(k-1)T - \theta, (k-1)T + \theta] \quad (4.24)$$

The constraints (4.24) require the departure of t_k to take place $(k-1)T$ minutes after the departure of train t_1 at every station, give or take θ minutes. This indeed results in one train leaving every T minutes after t_1 . As an example, in a time horizon of 4 hours, where $T = 60$ min, and a bandwidth of 5 minute, there are 4 cyclic trains yields the following periodic constraints according to (4.24):

$$\begin{aligned} x_2 - x_1 &\in [55, 65], \\ x_3 - x_1 &\in [115, 125], \\ x_4 - x_1 &\in [175, 185], \end{aligned}$$

However, in a feasible solution to the constraints (4.24), the starting time of two operations may lie apart more than an integer multiple of $T \pm \theta$ minutes. Consider two trains t_j and t_{j+1} , and suppose that $x_1 = 0$. Then in a feasible solution to (4.24), we may have $x_j = (j-1)T - \theta$ and $x_{j+1} = jT + \theta$. Thus, $x_{j+1} - x_j = 2\theta$. For the example above, a feasible solution would be $x_1 = 0$, $x_2 = 55$, $x_3 = 125$, $x_4 = 185$, which gives $x_3 - x_2 = 70$.

Therefore, the time of all other pair of train also need to be periodic.

Let us denote a set of *periodic activities* for every pair of periodic trains in exiting timetable as follows,

$$E_{peri} = \{(i, j) : i \in V^{ini} \text{ and } j \in V^{ini} \text{ are scheduled periodically}\}$$

then the periodic structure of exiting cyclic timetable can be restricted by,

Allowed deviations to the periodic structure:

$$x_j - x_i \leq (\pi_j - \pi_i) + \theta \quad \forall (i, j) \in E_{peri} \quad (4.25)$$

$$x_j - x_i \geq (\pi_j - \pi_i) - \theta \quad \forall (i, j) \in E_{peri} \quad (4.26)$$

We define the constrains for $i < j$ in order to prevent stating double constrains. Finally, for the four trains in the example, constraints (4.25-4.26) are the following:

$$\begin{aligned} x_2 - x_1 &\in [55, 65], \\ x_3 - x_1 &\in [115, 125], \\ x_4 - x_1 &\in [175, 185], \\ x_3 - x_2 &\in [55, 65], \\ x_4 - x_2 &\in [115, 125], \\ x_4 - x_3 &\in [55, 65]. \end{aligned}$$

4.2.5 Frequency of Additional Trains

For each train line, a frequency is specified, and a type, which determines the velocity and the stations that the line calls at. Suppose that the frequency of N additional trains t_1, \dots, t_N is to be synchronized, so that the departures of these trains are spread evenly across the considered time horizon T_{hor} . A train should depart every $T_{hor}/(N-1)$ minutes. Similar to the deviations of periodic structure defined in Section (4.2.4), we do not want to fix the insertion of the additional trains too much beforehand. In order to offer some flexibility in the departure times, a bandwidth β is defined, by which the departure time of a train may deviate from its perfect departure time. Then the frequency time window are defined as

$$x_i - x_1 \in \left[\frac{i-1}{N-1}T_{hor} - \beta, \frac{i-1}{N-1}T_{hor} + \beta \right]$$

Let us introduce a set of *frequency activities* for every pair of additional trains *within a train line*.

$$E_{syn} = \{(i, j) : i \in V^{add} \text{ and } j \in V^{add} \text{ are scheduled synchronously}\}$$

then the frequency constraints can be represented by,

Frequency for additional trains:

$$x_j \geq x_i \quad \forall (i, j) \in E_{syn}, i, j = 1, \dots, N \quad (4.27)$$

$$x_j - x_i \geq \frac{j-i}{N-1}T_{hor} - \beta \quad \forall (i, j) \in E_{syn}, i, j = 1, \dots, N \quad (4.28)$$

$$x_j - x_i \leq \frac{j-i}{N-1}T_{hor} + \beta \quad \forall (i, j) \in E_{syn}, i, j = 1, \dots, N \quad (4.29)$$

4.3 Objectives Functions

As analysed in Section 3.2, we consider objectives in the view of the following three aspects in this Chapter,

- high quality of the performance to the additional trains, which can be represented by the objective,

(1) minimizing travel time of additional trains

- low deviations to the initial services, which can be represented by the objective,

(2) minimizing the total adjustments to initial trains

- high quality of the entire new timetable, which can be represented by the objectives,

(3) minimizing the makespan of the new timetable,

(4) maximizing the robustness of the new timetable,

(5) minimizing the number of required train-sets,

For each of the five objectives, we define a function that expresses the objective in terms of the decision variables:

F_t for the average travel time of the additional trains,

F_a for the total adjustments of the initial trains,

F_m for the makespan of the new timetable,

F_r for the timetable robustness,

F_s for the number of required train-sets.

Each of these functions can be substituted for the general objective function $F(x)$, where x is the vector of variables. One can also assign weights to the five functions above, and use a weighted multi-objective function.

The following sections formulate the four functions F_t , F_a , F_m and F_r . For the train-set function F_s , we require some ideas that are introduced in Chapter 5. Therefore, the train-sets function is not defined until Chapter 5.

It is clearly that the objective functions F_t , F_a and F_m can be formulated respectively as follows,

$$F_t = \sum_{t \in \mathcal{T}^{add}} (x_{last_t} - x_{first_t})/N \quad (4.30)$$

$$F_a = \sum_{i \in V^{ini}} |x_i - \pi_i| \quad (4.31)$$

$$F_m = x_{end} - x_{begin} \quad (4.32)$$

here, $first_t$ and $last_t$ are the first and last event of train t respectively. $begin$ and end are the beginning and ending event of the entire timetable. Their time instance can be achieved by $x_{begin} = \min x_i$ and $x_{end} = \max x_i$ respectively, where $\forall i \in V$. In addition, function (4.31) can be rewrote as follows by introducing an auxiliary variable ad_i ,

$$x_i - \pi_i \leq ad_i \quad i \in V^{ini} \quad (4.33)$$

$$\pi_i - x_i \leq ad_i \quad i \in V^{ini} \quad (4.34)$$

then, $F_a = \sum_{i \in V^{ini}} ad_i$.

Next, we will describe the function of robustness in the ATP problem in more detail.

The timetable robustness can be improved by pulling apart trains that share a track. If there is a lot of times between two consecutive trains, these times can be used as buffers in case of delays. Leon (2003) modelled an robustness cyclic timetable by setting the interval of trains be closed to the middle of the time window in order to pull apart each other. For the robustness in the ATP problem, a trade-off has to be made, however, between increasing the interval time between trains on one hand, and decreasing the modifications to initial timetable on the other hand. For example, during inserting additional trains, although the involved trains share the track for entering or leaving the station, the requirement of minimizing deviations to initial timetable implies that these trains can not be pulled apart too far.

Then the objective of robustness in the ATP problem restraints that a additional train should be inserted in the position that

- (1) between two trains that of largest idle interval, and simultaneously,
- (2) on or be closed to the middle of the interval time.

An example of a particular kind of this situation is shown in Figure 4.6. Three existing trains t_1, t_2 and t_3 are scheduled in the initial timetable, and a new train needs to be inserted. Both arrival and departure headway are set to be 3 min. The additional train could be inserted between t_2 and t_3 as train path t_n shown in Figure 4.6a, or t_1 and t_2 as train path t'_n shown in Figure 4.6b. Both of the solutions do not lead any deviation to initial services. However, the robustness of the new timetables are completely different, in practice timetable in Figure 4.6b is preferred due to a better robustness. Merely minimizing the modifications or trip times does not suit our goal to get a robustness insertion.

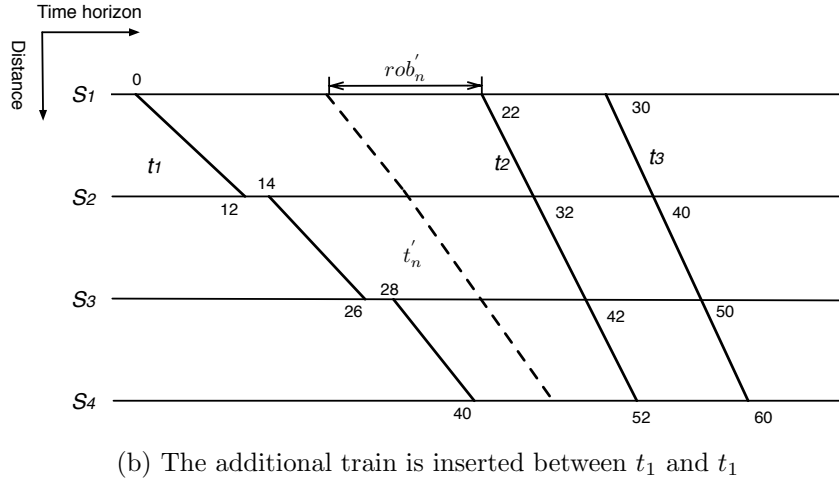
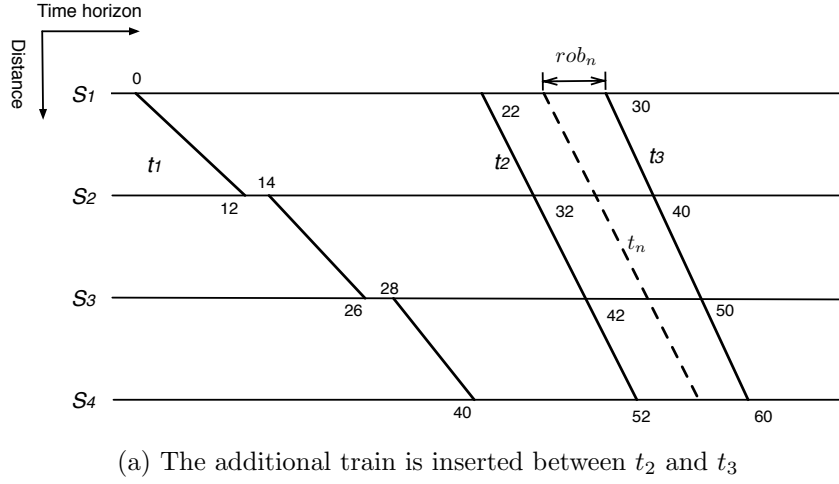


Figure 4.6: Different robustness results from different insertion

Let $E_{headway}^r = (v, u)$ be the set of headway activities corresponding to the safety constraints, where $v \in V^{add}, u \in V$. Then, the departure times of trains are pulled apart

when the process time for $e \in E_{headway}^r$ is increased, i.e. the additional trains are inserted in the middle of the adjacent trains between which there has the largest time interval.

Therefore, introduce an auxiliary variable γ_e for all $e \in E_{headway}^r$. The auxiliary variable γ_e is constrained as follows.

$$\gamma_e \geq x_i - x_j \quad \forall e = (i, j) \in E_{headway}^r \quad (4.35)$$

$$\gamma_e \geq x_j - x_i \quad \forall e = (i, j) \in E_{headway}^r \quad (4.36)$$

So $\gamma_e \geq |x_i - x_j|$. Then define the parameter rob_i as

$$rob_i = \min \gamma_e \quad \forall e = (i, j) \in E_{headway}^r \quad (4.37)$$

That is, rob_i denotes the minimum time interval between the additional activity i and other activities. Thus, maximizing rob_i means pushing x_i away from the other trains, and thus insert the additional trains in the middle of largest time interval.

Using the above, the robustness objective function is defined as

$$F_r = \sum_{i \in V^{add}} rob_i \quad (4.38)$$

Recall that the function F_r is maximized. This ensures that the additional trains are inserted in the middle of largest time intervals. In other words, maximizing F_r means maximizing the new timetable robustness.

Note that as there is a difference between the speed of trains, not only a single solution is computed, but a set of optimal solutions with respect to travel time and the robustness objective, as illustrated in Figure 4.7. Thus the solutions give a clear indication of where the train should be added when the objective of robustness is taken into account.

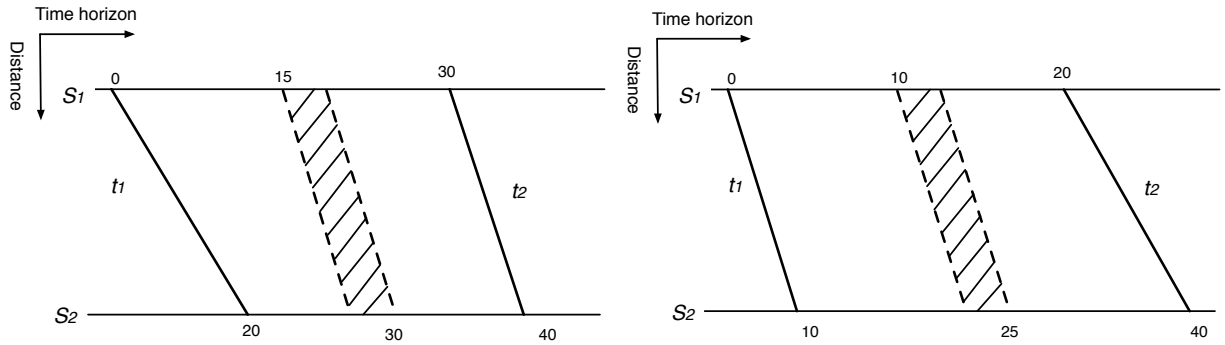


Figure 4.7: Optimal train paths for a request on maximum robustness

4.4 Experimental Studies

4.4.1 Purpose

The experiments we have conducted are in two parts. The first part investigates the influence of using different objective functions and different tolerance of frequency on

the values of identified performance measures, while the second part analyses if and how various tolerance of disruptions to initial timetable influences the insertion effect.

4.4.2 Scenarios and Experiment Setting

The railway traffic in the Shanghai-Hangzhou HSR, China has been used to generate the scenarios. This rail line consists of double-tracked HSR lines that are the major links station Shanghai Hongqiao (SHHQ), Songjian South (SJS), Jinshan North (JSN), Jiashan South (JSS), Jiaxing South (JXS), Tongxiang (TX), Haining West (HNW), Yuhang South (YHS) as well as Hangzhou (HZ). The cyclic nature of the timetable is illustrated in Figure 4.8.

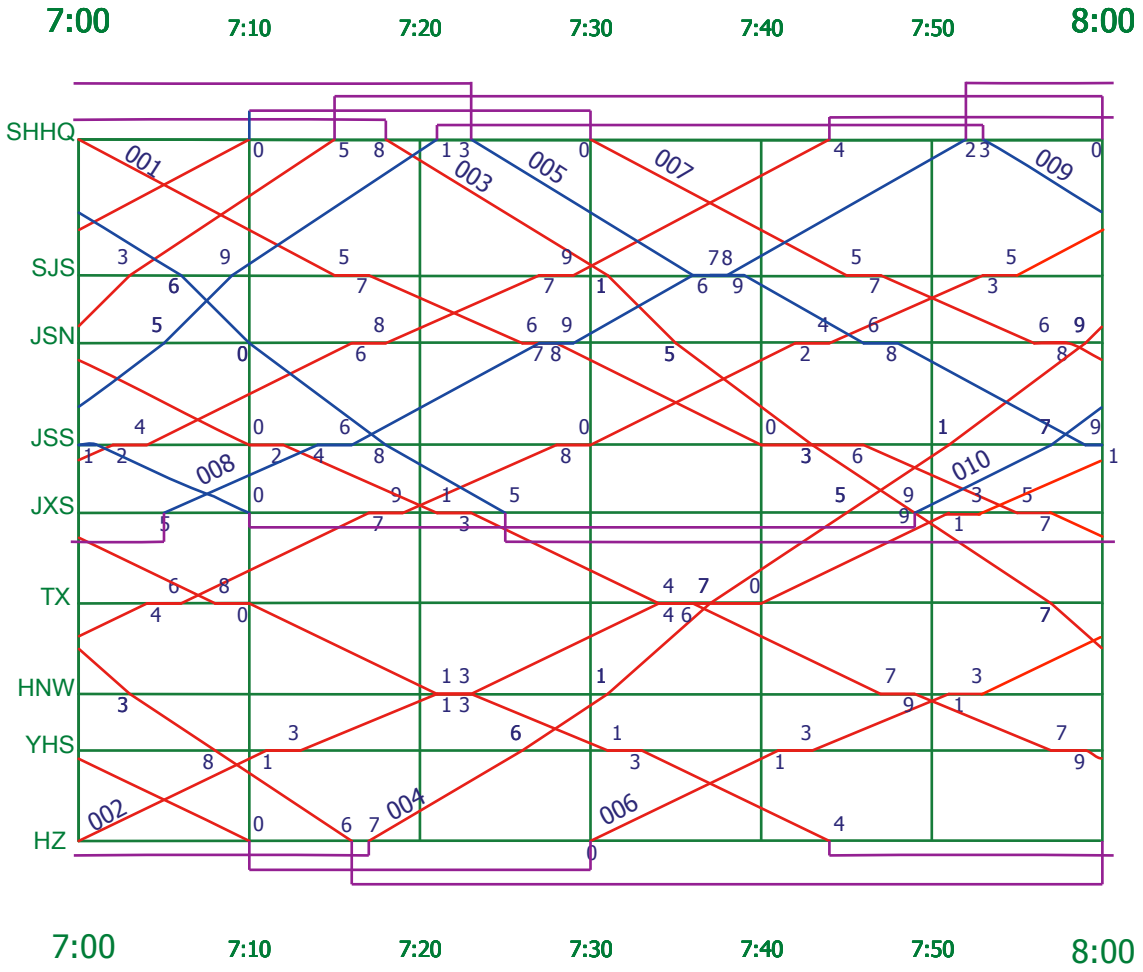


Figure 4.8: One hour time-space diagram for the track between Shanghai and Hangzhou

Two different time horizon of the initial timetable have been used for the computational experiments. The timetable of different time horizon are:

- (a) Timetable (5:00-14:00): this is a timetable with the time horizon from 5:00 to 14:00. In this case, there are in total 79 trains in both directions.

- (b) Timetable (5:00-23:00): this is a timetable with the time horizon from 5:00 to 23:00. The traffic data includes 159 trains.

In the experiments, minimum headways are set to 3 minutes for both consecutive arrivals and departures. Acceleration and deceleration times are set to 2 and 1 minutes respectively for both high-speed and medium-speed trains. In addition, taking the variable velocity into consideration, maximum driving time is set to 110% (w.r.t. the minimum driving time). All of the transfer connections included are considered non-breakable and take place at the station JXS. The minimum required connection time is 3 min but in the timetable, there are often between 20 and 50 min in between incoming and connecting trains.

We have used a number of objective functions and performance measures, based on discussions with people from the railway industry and according to the discussion in Section 3.2. The objectives we have considered (each separately) in this section are:

- (A) Maximise the robustness F_r to the new timetable.
- (B) Minimise the makespan F_m to the new timetable.
- (C) Minimise the trip time F_t to additional trains.
- (D) Minimise the total adjustments F_a to initial timetable.

The performance measures, or rather solution quality measures, we have considered independent of objective function are:

- Robustness of the insertion.
- Makespan of the new timetable.
- Average trip time of the additional trains.
- Total modifications of the initial timetable.

The robustness of timetable here is reflected and calculated by equation 4.38. Makespan refers to the total consumed time for all of the considered trains. Average trip time refers to the average time consumed for an additional train from its original departure to final destination, including the running times and dwell times. Total modifications refers to the sum of all adjustments of the all positive adjustments of the initial trains at each visiting stations. This also can be rewrote to total cost modifications which consists of one part that is train-related and associated with the punishment of adjustments for each train.

In order to properly test the solution techniques a variety of test problems must be generated. Choosing a set of unbiased test problems however is not simple. For example

the number of current trains must be decided as should the number of new trains. For new trains, the various tolerance of frequency β should be evaluated. For initial trains, operations may be fixed or tolerance of disruptions including allowable adjustments Δ and periodic structure θ may be set. Furthermore the size of the disruption tolerance can be set in many different ways.

With these issues in mind, we have decided to apply the techniques to some of the more important parts of the ATP problem process. Firstly (Part 1) the existing schedules is fixed and the insertion of additional services subject to various level of frequency β is tested. Secondly (Part 2) an existing schedule is taken and additional train services with fixed β are inserted. Existing services are given various level of disruption tolerance and may be shifted if improvements to the optimal objectives can be realised.

We have programmed a adding train paths tool in Visual Basic language as shown in Appendix C, and solved the ATP problems on a PC with Intel(R) Core(TM) i3-2120 CPU @ 3.30GHZ + 3.30GHZ and 4 GB of RAM, and all the algorithms and examples in this section are implemented in Visual Studio 2013 on the Windows 7, 64 bit. IBM ILOG Cplex 12.5 with default set is used as a solver.

4.4.3 Result

Due to the extent of the numerical investigations, the results have been condensed greatly and only a summary of the most important results are shown.

Part 1: Objectives and quality measures with various tolerance of frequency constraint

In this part of the numerical investigations, existing trains are *fixed* and various frequency β is set.

We have applied the objective functions presented in Section 4.3 and computed the values of the performance measures for the generated insertion solutions. The purpose is to investigate how a certain objective function represents the overall (collection of) associated performance measures and how the various frequency constraints β effect.

We solved the ATP problem for inserting 10-20 additional trains and with various frequency constraints of $\beta = 10, 15$ and 20 using objective function (A)-(C). In Figures (4.9-4.11), the results with the initial timetable (5:00-14:00) from using objective function (A), (B) and (C) respectively can be seen.

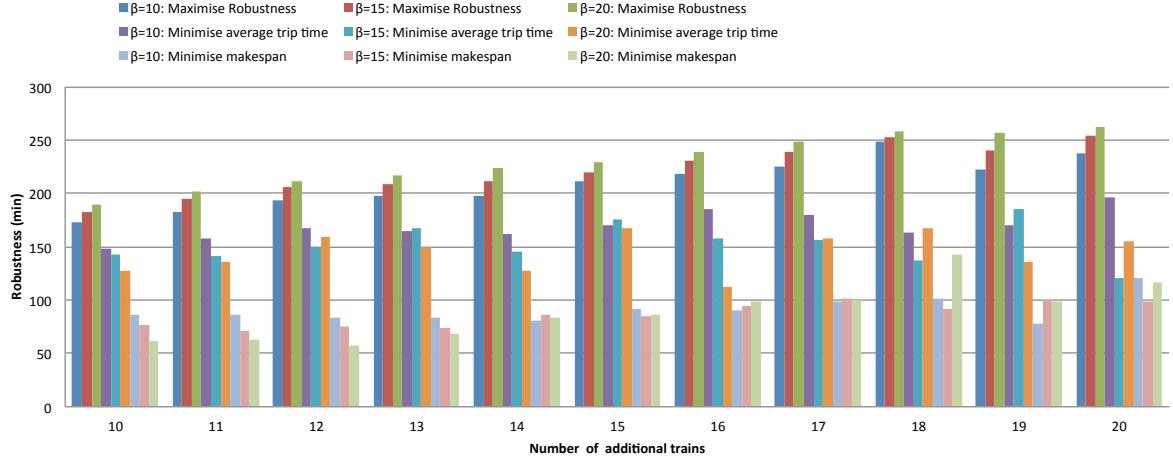


Figure 4.9: Overview the values of the performance measure *robustness of the insertion* based on objective function applied and number of additional trains for each scenario

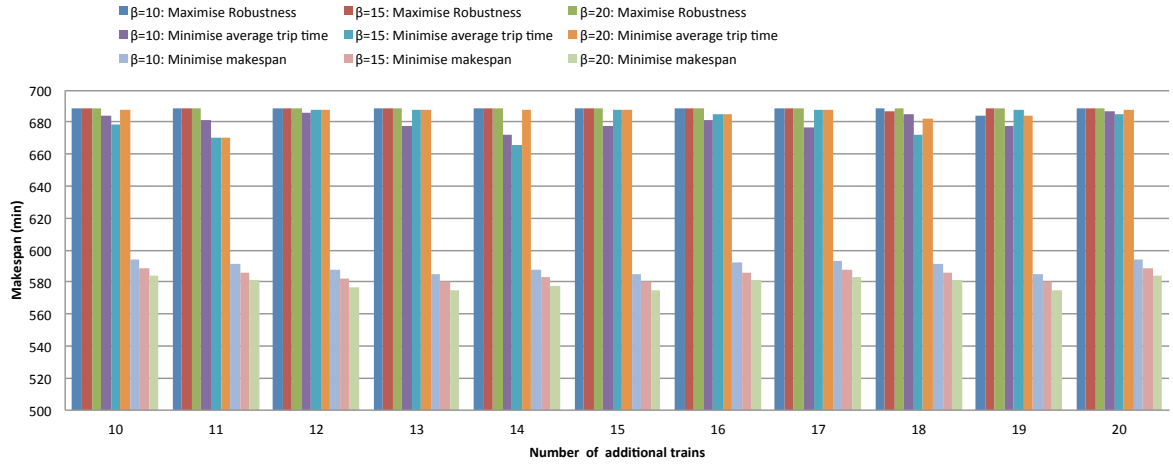


Figure 4.10: Overview the values of the performance measure *makespan of the new timetable* based on objective function applied and number of additional trains for each scenario

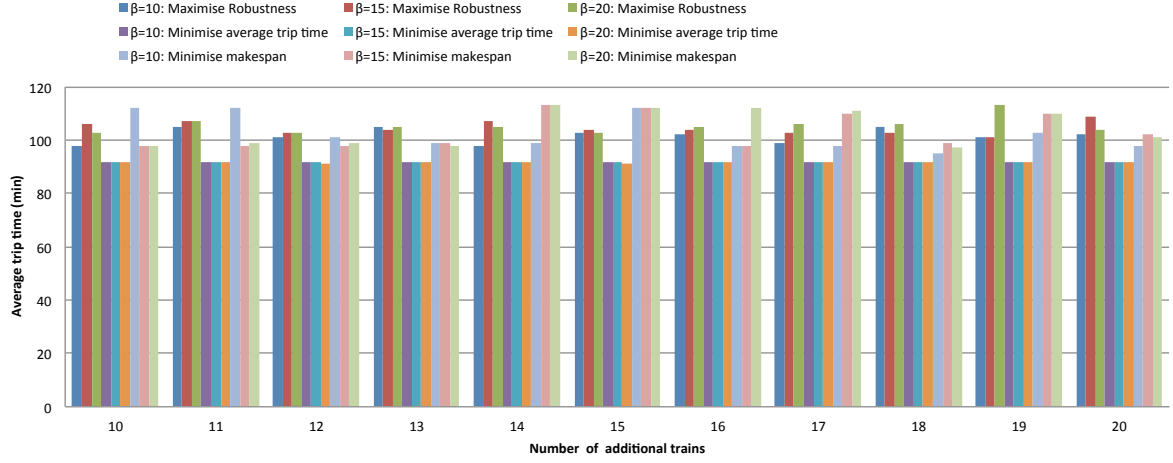


Figure 4.11: Overview the values of the performance measure *average trip time of additional trains* based on objective function applied and number of additional trains for each scenario

It is obviously that the higher tolerance of frequency constraints (the bigger β), implies the more freedom for trains insertion, and the better objectives can be achieved including bigger robustness, smaller makespan and average trip time. More importantly, different objectives have different sensitivity to the frequency constraints, such as the objective of trip time has the lowest sensitivity (as shown in Figure 4.11), while the objective of robustness has the highest sensitivity (as shown in Figure 4.9) to the different tolerance of frequency constraints. Figure 4.12 demonstrates the different inserting solutions with different tolerance of frequency to the objective of maximising robustness. Better robustness can be achieved with more freedom of insertion which is implied by bigger β here.

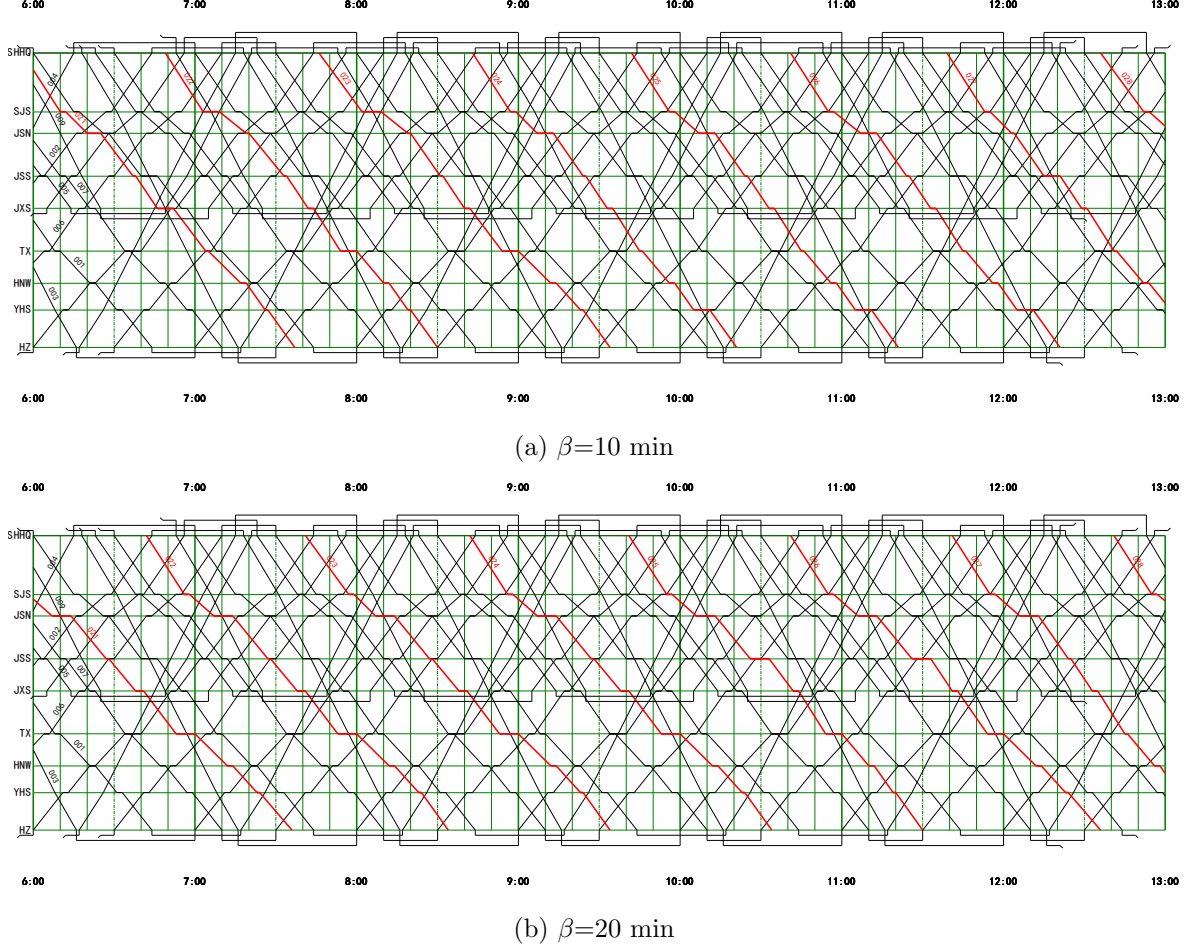


Figure 4.12: Different insertions to maximise robustness based on different tolerance of frequency constrains $\beta = 10$ or $\beta = 20$.

Figure 4.10 shows that a minimising of the average trip time or maximising robustness tend to a very high makespan. This can be explained by the fact that the capacity utilization is much lower in the early and nigh time because of the maintenance time. The the additional trains tend to be inserted in these time periods to generate a new timetable with a low trip time or a high robustness, which however will cause a high makespan. Whereas, due to the restricted dwell time at stations and running time in sections, no big difference on the performance measure of average trip time even though using various objective functions, which is show in Figure 4.11.

The scenarios were also solved with the initial timetable (5:00-23:00) and the same tendencies were seen.

The CPU times required typically increased with the level of frequency β , and also be different in the various objective functions. Table 4.4 shows the CPU times of the above instances. For the objective function (B), minimising the makespan of the new timetable, the CPU times are negligible, i.e. no more than several seconds.

Table 4.4: CPU times (in seconds) for inserting 10-20 trains

	Nr. of \mathcal{T}^{add}	$\beta = 10$ min			$\beta = 15$ min			$\beta = 20$ min		
		(A) ¹	(B) ²	(C) ³	(A) ¹	(B) ²	(C) ³	(A) ¹	(B) ²	(C) ³
(a) Time horizon:	10	6	1	1	40	0	1	67	0	4
5:00-14:00	11	6	1	1	17	0	3	53	1	10
	12	30	1	1	45	1	3	35	1	14
	13	38	1	1	25	1	4	53	1	28
	14	16	1	2	41	1	12	58	1	51
	15	56	1	2	41	1	22	109	1	77
	16	76	1	3	45	1	15	88	2	97
	17	22	1	3	46	1	27	245	2	230
	18	164	2	9	155	2	101	1055	3	1767
	19	41	2	5	322	2	61	268	3	441
	20	31	2	5	194	2	40	571	3	644
(b) Time horizon:	10	7	1	1	26	1	2	50	1	3
5:00-23:00	11	9	1	2	39	1	3	23	1	5
	12	17	1	2	53	1	4	119	1	11
	13	42	1	2	62	1	4	185	1	19
	14	38	2	3	109	2	4	185	2	22
	15	81	2	4	40	2	17	1003	2	35
	16	115	2	5	209	2	18	123	2	25
	17	195	2	4	149	2	9	2261	3	20
	18	150	2	5	196	2	18	954	3	45
	19	115	2	5	251	3	31	759	3	147
	20	124	3	9	197	3	57	1550	3	253

¹ Using the objective function (A): maximise the robustness to the new timetable.

² Using the objective function (B): minimise the makespan to the new timetable.

³ Using the objective function (C): minimise the trip time to additional trains.

Part 2: Objectives with various disruptions

In this part of the numerical investigations, initial trains are *unfixed* and various tolerance of disruptions to initial timetable which consists of Δ and θ are set.

For the experiments, we inserted 10 and 20 additional trains with a fixed frequency $\beta = 10$. The existing timetable may be adjusted under the different tolerance of disruptions as described in Section 4.2.4. Table 4.5 shows a complete description of the instances constructed using various allowed adjustment Δ and periodic structure θ . Also, these 20 instances are successively assessed with the four objective functions described above.

Table 4.5: Set of instances with different tolerance of disruptions

Instance Nr.	Tolerance		Instance Nr.	Tolerance	
	Δ (min)	θ (min)		Δ (min)	θ (min)
1	1	0	11	4	1
2	1	1	12	4	2

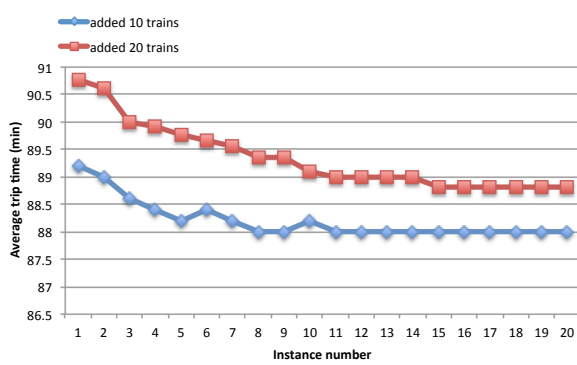
(Continued on next page)

Table 4.5 – (continued from previous page)

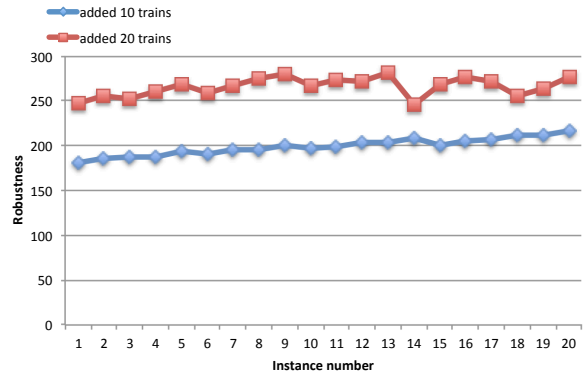
Instance Nr.	Tolerance		Instance Nr.	Tolerance	
	Δ (min)	θ (min)		Δ (min)	θ (min)
3	2	0	13	4	3
4	2	1	14	4	4
5	2	2	15	5	0
6	3	0	16	5	1
7	3	1	17	5	2
8	3	2	18	5	3
9	3	3	19	5	4
10	4	0	20	5	5

As shown in Part 1, all of the 20 additional trains can be inserted to the fixed initial timetable, then when the objective function of (D) minimizing total adjustments is applied, the total adjustments will always equal to zero even the initial timetable is allowed to be rescheduled. With the same β , when inserted with the objective function of (B) minimizing the makespan, the objectives are similar in 20 instances.

Next, we will discuss in detail the results with objective function (A) and (C). Figure 4.13 shows the trends of using the other two different objective functions of (C) minimizing trip time and (A) maximizing robustness in these 20 instances to insert 10 or 20 trains in the initial timetable (5:00-14:00). Figure 4.13a implies that as the tolerance of disruptions increases, the average trip time decreases and becomes flat finally. However, the robustness of insertions increase gently as shown in Figure 4.13b. The reason for the trend is that the objective function (A) describe by formula (4.38) evaluates the robustness of insertion position, consequently the initial trains are pulled apart from the new trains to gain robustness.



(a) Trends of average trip time using objective function (C) in various instances



(b) Trends of robustness using objective function (A) in various instances

Figure 4.13: Trends of the objectives based on the initial timetable (5:00-14:00)

All of trains can be inserted without leading any deviations to the initial trains, however, when the initial timetable could be rescheduled, there is a *trade-off* between

the minimum total adjustments and minimum trip time, and also between the minimum total adjustments and maximum robustness.

To evaluate the impact of different objectives on the total adjustments, four scenarios of instances, each one emphasizing one aspect of the objective function, have been considered. The scenarios are:

- Scenario 1: emphasize maximization of robustness: objective function (A) is used and the strongly penalize the robustness.
- Scenario 2: balance adjustments and maximizing of robustness: an equilibrated combination of coefficients in the objective function (A) and (D) is considered to maximize

$$w_r F_r - w_a F_a$$

The coefficients are $w_r = 20$, $w_a = 1$. These values were selected to permit some decrease of robustness if adjustments become too large.

- Scenario 3: emphasize minimization of average trip time: objective function (C) is used and the strongly penalize the trip time.
- Scenario 4: balance adjustments and minimizing of trip time: an equilibrated combination of coefficients in the objective function (C) and (D) is considered to minimize

$$w_t F_t + w_a F_a$$

The coefficients are $w_t = 30$, $w_a = 1$. These values were selected to permit some increase of trip time if adjustments become too large.

In Table 4.6, comparisons of the result for each instance are reported. Column (Total adjustments (%)) is total adjustments in minutes to the initial timetable. The arrows and numbers in parenthesis is the relative changes between the emphasizing objective function and its corresponding balance objective function, such as scenario 1 and 2, and scenario 3 and 4. The relative changes is computed as follows: $GAP_y = (y - y^*)/y^*$. Column (Robustness/Trip time (%)) is the robustness for scenario 1 and 2, and average trip time for scenario 3 and 4, respectively. The arrows and numbers in parenthesis is the relative changes too.

The result of Table 4.6 highlights the impact of the different objective functions on the total adjustments when the initial timetable is unfixed. This table presents the optimal results depending on the type of objectives the emphasis is put on. On one hand, by introducing the balance objectives, the value of robustness and trip time is not significantly affected, but the value of total adjustments is decreased dramatically. This implies that a balance objective which takes the adjustments into account in more

appropriate in practice. On the other hand, comparing the results of timetable (5:00-23:00) with timetable (5:00-14:00), in scenario 1 and 3 the values of total adjustments are much larger in the timetable (5:00-23:00), which implies that more initial trains are affected in the timetable (5:00-23:00) even though inserting the same number of additional trains. This may be explained by the constraints of periodic structure; more number of cycle time are involved.

The result also show that the instances of timetable (5:00-23:00) are harder to solve than instances of timetable (5:00-14:00). Such results can be explained by a greater number of events to be rescheduled in timetable (5:00-23:00). Moreover, the higher level of disruption tolerance to initial timetable also lead to more CPU times. This is caused by more chances of insertions and adjustments. In addition, achieving the balance objectives is more time consumed than the emphasizing objectives. However, the consumed times in all of the instances are acceptable for a tactical or short-term planning.

4.4.4 Summary of the Computational Results

In this section, we report on the performance of the models for inserting additional trains to an existing cyclic timetable.

The numerical investigation consists of two parts. In Part 1, the existing cyclic timetable is fixed. The influence of using various tolerance of frequency constraints β and different objective functions is investigated in this part. In Part 2, the existing cyclic timetable is unfixed. This part analyses if and how various tolerance of disruptions to initial timetable influences the insertion effect.

Part 1 compares the performances with various β and objective functions. (i) The performance of trip time changes a little with different β and objectives due to the constraints of running and dwell time. (ii) The CPU time changes a lot, depending on the setting of β and objective functions. The bigger β indicates more freedom for insertion and consequently increases the required computational time. Overall, we are able to insert 10-20 trains with minimum travel time within 3 seconds. It is the hardest to solve the ATP problem with maximise the robustness, however, the computational time is also within 40 min, depending on a fixed existing timetable. (iii) The change of β has different effect on different objective function. The objective of maximum robustness is susceptible on various β . This can be explained by the fact that the additional trains have more freedom to get a robustness insertion with a higher tolerance of frequency constraint.

Part 2 compares the performances with various tolerance of disruptions, including allowed adjustments and deviations to periodic structure. (i) With the increase of disruption tolerance, the trip time decreases and becomes flat finally, while the robustness rises smoothly. (ii) There is a trade-off between total adjustments and other objectives. By applying a balance objective, the trip time and robustness are close to that with the

emphasize objectives, and more importantly, the total adjustments can be substantially decreased. A multi-objective function is appropriate to get a better insertion, because the total adjustments becomes a essentially important criteria when the initial timetable is unfixed.

Table 4.6: Result for the instances of inserting 10 trains

		Time horizon: 5:00-14:00				Time horizon: 5:00-23:00			
		Objective	Total	Robustness ¹ /	CPU time	Objective	Total	Robustness ¹ /	CPU time
		Nr.	value	adjustments(%) ²	Trip time ¹ (%) ²	(s)	value	adjustments(%) ²	Trip time ¹ (%) ²
(1) Emphasize maximization of robustness	1	181	520	181	11	136	996	136	37
	2	185	481	185	25	141	1026	141	34
	3	187	814	187	22	137	1733	137	49
	4	188	901	188	39	143	1862	143	101
	5	193	892	193	38	148	1801	148	127
	6	190	1169	190	26	141	2552	141	40
	7	195	1246	195	44	146	2392	146	76
	8	196	1217	196	68	152	2462	152	103
	9	201	1228	201	37	157	2412	157	79
	10	197	1557	197	42	144	2928	144	58
	11	198	1509	198	24	150	3125	150	89
	12	203	1470	203	47	155	2726	155	183
	13	204	1580	204	42	161	3015	161	117
	14	209	1923	209	42	166	2877	166	439
	15	200	2065	200	46	146	3183	146	40
	16	205	1744	205	26	153	3272	153	175
	17	206	1698	206	70	158	3740	158	136
	18	211	2447	211	42	164	3207	164	238
	19	212	1856	212	84	169	3119	169	316
	20	217	2279	217	52	174	3988	174	101
(2) Balance adjustments and maximizing of robustness	1	3556	64 (↓ 712.5)	181 (↓ 0.0)	51	2648	72 (↓ 1283.3)	136 (↓ 0.0)	62
	2	3660	40 (↓ 1102.5)	185 (↓ 0.0)	40	2790	30 (↓ 3320.0)	141 (↓ 0.0)	68
	3	3612	108 (↓ 653.7)	186 (↓ 0.5)	41	2648	72 (↓ 2306.9)	136 (↓ 0.7)	54
	4	3684	76 (↓ 1085.5)	188 (↓ 0.0)	69	2790	30 (↓ 6106.7)	141 (↓ 1.4)	200

(Continued on next page)

Table 4.6 – (continued from previous page)

Time horizon: 5:00-14:00						Time horizon: 5:00-23:00			
Instance	Objective	Total	Robustness ¹ /	CPU time		Objective	Total	Robustness ¹ /	CPU time
Nr.	value	adjustments(%) ²	Trip time ¹ (%) ²	(s)		value	adjustments(%) ²	Trip time ¹ (%) ²	(s)
5	3779	81 (↓ 1001.2)	193 (↓ 0.0)	56		2903	57 (↓ 3059.6)	148 (↓ 0.0)	132
6	3612	108 (↓ 982.4)	186 (↓ 2.1)	40		2648	72 (↓ 3444.4)	136 (↓ 3.5)	109
7	3718	142 (↓ 777.5)	193 (↓ 1.0)	76		2790	30 (↓ 7873.3)	141 (↓ 3.4)	197
8	3788	132 (↓ 822.0)	196 (↓ 0.0)	70		2903	57 (↓ 4219.3)	148 (↓ 2.6)	193
9	3851	149 (↓ 724.2)	200 (↓ 0.5)	79		3003	137 (↓ 1660.6)	157 (↓ 0.0)	194
10	3612	108 (↓ 1341.7)	186 (↓ 5.6)	49		2648	72 (↓ 3966.7)	136 (↓ 5.6)	139
11	3718	162 (↓ 831.5)	194 (↓ 2.0)	75		2790	30 (↓ 10316.7)	141 (↓ 6.0)	304
12	3788	132 (↓ 1013.6)	196 (↓ 3.4)	120		2903	57 (↓ 4682.5)	148 (↓ 4.5)	129
13	3859	221 (↓ 614.9)	204 (↓ 0.0)	89		3003	137 (↓ 2100.7)	157 (↓ 2.5)	282
14	3926	174 (↓ 1005.2)	205 (↓ 1.9)	128		3090	170 (↓ 1592.4)	163 (↓ 1.8)	292
15	3612	108 (↓ 1812.0)	186 (↓ 7.0)	93		2648	72 (↓ 4320.8)	136 (↓ 6.8)	102
16	3718	142 (↓ 1128.2)	193 (↓ 5.9)	104		2790	30 (↓ 10806.7)	141 (↓ 7.8)	308
17	3788	132 (↓ 1186.4)	196 (↓ 4.9)	82		2903	57 (↓ 6461.4)	148 (↓ 6.3)	234
18	3859	221 (↓ 1007.2)	204 (↓ 3.3)	155		3003	137 (↓ 2240.9)	157 (↓ 4.3)	587
19	3930	190 (↓ 876.8)	206 (↓ 2.8)	106		3090	170 (↓ 1734.7)	163 (↓ 3.6)	234
20	3975	245 (↓ 830.2)	211 (↓ 2.8)	100		3162	238 (↓ 1575.6)	170 (↓ 2.3)	503
(3) Emphasize minimization of average trip time	1	89	512	89	1	91	1058	91	6
	2	89	522	89	1	90	1049	90	6
	3	89	1004	89	1	90	2082	90	12
	4	88	1054	88	1	90	1902	90	16
	5	88	1064	88	1	90	2018	90	14
	6	88	1465	88	1	90	2657	90	14
	7	88	1477	88	1	89	3036	89	21
	8	88	1603	88	1	89	2777	89	20
	9	88	1592	88	1	89	3003	89	25

(Continued on next page)

Table 4.6 – (continued from previous page)

		Time horizon: 5:00-14:00				Time horizon: 5:00-23:00			
Instance		Objective	Total	Robustness ¹ /	CPU time	Objective	Total	Robustness ¹ /	CPU time
Nr.		value	adjustments(%) ²	Trip time ¹ (%) ²	(s)	value	adjustments(%) ²	Trip time ¹ (%) ²	(s)
	10	88	1794	88	1	89	3674	89	12
	11	88	1905	88	1	89	3919	89	12
	12	88	2065	88	1	89	4173	89	13
	13	88	2093	88	1	89	4418	89	13
	14	88	2125	88	1	89	4371	89	13
	15	88	2246	88	1	89	4747	89	19
	16	88	2180	88	1	89	5078	89	30
	17	88	2262	88	1	89	4574	89	20
	18	88	2688	88	1	89	4697	89	36
	19	88	2616	88	1	89	5549	89	27
	20	88	2662	88	1	89	5242	89	43
(4) Balance adjustments and minimizing of trip time	1	2698	22 (↓ 2227.3)	89 (↓ 0.0)	2	2750	14 (↓ 7457.1)	91 (↑ 0.0)	21
	2	2688	18 (↓ 2800.0)	89 (↓ 0.0)	2	2730	24 (↓ 4270.8)	90 (↑ 0.0)	14
	3	2698	22 (↓ 4463.6)	89 (↓ 0.0)	5	2750	14 (↓ 14771.4)	91 (↑ 1.3)	31
	4	2688	18 (↓ 5755.6)	89 (↓ 0.7)	5	2730	24 (↓ 7825.0)	90 (↑ 0.0)	31
	5	2681	20 (↓ 5220.0)	89 (↓ 0.9)	6	2726	26 (↓ 7661.5)	90 (↑ 0.0)	30
	6	2698	22 (↓ 6559.1)	89 (↓ 0.7)	6	2750	14 (↓ 18878.6)	91 (↑ 1.7)	39
	7	2688	18 (↓ 8105.6)	89 (↓ 0.9)	5	2730	24 (↓ 12550.0)	90 (↑ 0.7)	45
	8	2681	20 (↓ 7915.0)	89 (↓ 1.1)	6	2726	26 (↓ 10580.8)	90 (↑ 0.7)	40
	9	2681	20 (↓ 7860.0)	89 (↓ 1.1)	7	2726	26 (↓ 11450.0)	90 (↑ 0.7)	61
	10	2698	22 (↓ 8054.5)	89 (↓ 0.9)	8	2750	14 (↓ 26142.9)	91 (↑ 2.1)	44
	11	2688	18 (↓ 10483.3)	89 (↓ 1.1)	6	2730	24 (↓ 16229.2)	90 (↑ 1.4)	50
	12	2681	20 (↓ 10225.0)	89 (↓ 1.1)	6	2726	26 (↓ 15950.0)	90 (↑ 1.4)	57
	13	2681	20 (↓ 10365.0)	89 (↓ 1.1)	5	2726	26 (↓ 16892.3)	90 (↑ 1.4)	89
	14	2681	20 (↓ 10525.0)	89 (↓ 1.1)	5	2726	26 (↓ 16711.5)	90 (↑ 1.4)	78

(Continued on next page)

Table 4.6 – (continued from previous page)

Instance Nr.	Time horizon: 5:00-14:00					Time horizon: 5:00-23:00				
	Objective	Total	Robustness ¹ /	CPU time		Objective	Total	Robustness ¹ /	CPU time	
	value	adjustments(%) ²	Trip time ¹ (%) ²	(s)		value	adjustments(%) ²	Trip time ¹ (%) ²	(s)	
15	2698	22 (↓ 10109.1)	89 (↓ 1.1)	9		2750	14 (↓ 33807.1)	91 (↑ 2.5)	72	
16	2688	18 (↓ 12011.1)	89 (↓ 1.1)	8		2730	24 (↓ 21058.3)	90 (↑ 1.4)	61	
17	2681	20 (↓ 11210.0)	89 (↓ 1.1)	10		2726	26 (↓ 17492.3)	90 (↑ 1.4)	75	
18	2681	20 (↓ 13340.0)	89 (↓ 1.1)	7		2726	26 (↓ 17965.4)	90 (↑ 1.4)	96	
19	2681	20 (↓ 12980.0)	89 (↓ 1.1)	9		2726	26 (↓ 21242.3)	90 (↑ 1.4)	102	
20	2681	20 (↓ 13210.0)	89 (↓ 1.1)	7		2726	26 (↓ 20061.5)	90 (↑ 1.4)	125	

¹ The value of robustness for scenario 1 and 2, and value of average trip time for scenario 3 and 4.

² $GAP_y = (y - y^*)/y^*$; y^* , better solution value.

Chapter 5

Integration of Train-set Circulation and Adding Train Paths Problem

In this chapter, our main contribution is that we integrate the train set circulation to the ATP model described in Chapter 4 in order to decide simultaneously additional trains' schedules, initial trains' adjustments and train-set circulation for the new timetable.

Section 5.1 introduces the various use of train-set in China's HSR. Section 5.2 describes the basic terminologies and gives a literature overview on train-set circulation problem. Next, Section 5.3 describes special features of train-set planning in the ATP problem and the train-set circulation is decomposed into two sub-problem that keeps the scheduled train-set circulation to the initial trains and simultaneously covers the additional trains with minimum number of train-sets. We model the train-set circulation in ATP problem in Section 5.4. In order to solve the problem in a reasonable time, we start from fixed train-set route, and then apply flexible train-set route. In this process, we deal with the over-night turn-arounds and provide them possible alternative turning activities to decrease the waiting time of a train-set. Finally, the proposed model is tested in Section 5.5 based on Shanghai-Hangzhou HSR line in China. Several affecting factors are evaluated such as the various level of tolerance of disruptions, different objective functions, the introduction of time window constraints and the different use of train-sets.

5.1 Use of Train-set in China's HSR

In China's HSR, there exist two different approaches on the use of train-set according to the operating area of a train-set.

5.1.1 Fixed Use of Train-set

The first approach is called *fixed use* for short, which refers to that a train-set runs only among *certain operating sections*. It is consistent with the application of rolling stocks in conventional railway lines.

Figure 5.1 illustrates the fixed use of a train-set on a single section that the train-set first carries out the journey from terminus s_1 to terminus s_2 , and subsequently a reverse journey from s_2 to s_1 .

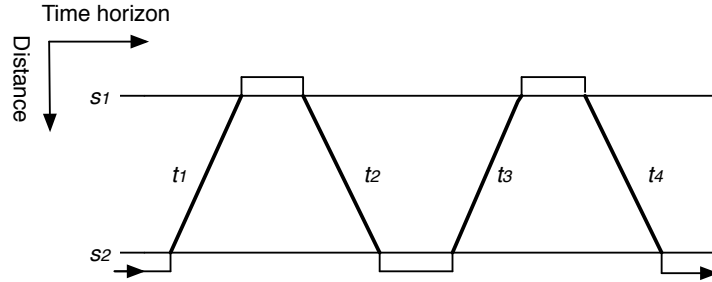


Figure 5.1: Fixed use of train-set

As the operating sections are certain, the fixed use of train-set is conducive to scheduling the train-set circulation. It is also convenient for the transportation maintenance. When disturbance occur, the influenced area is relative small and it is much more easy to manage the disruption. However, the fixed use also has lots of disadvantages, such as the low efficiency of train-set may caused by long waiting times at the turnaround station.

5.1.2 Flexible Use of Train-set

The second approach is called *flexible use* for short, which refers to that a train-set runs on *uncertain operating sections*. Instead of fixing operating sections, in the flexible use a train-set can carry trains at any sections on the condition that the requirement of turning around connection is satisfied.

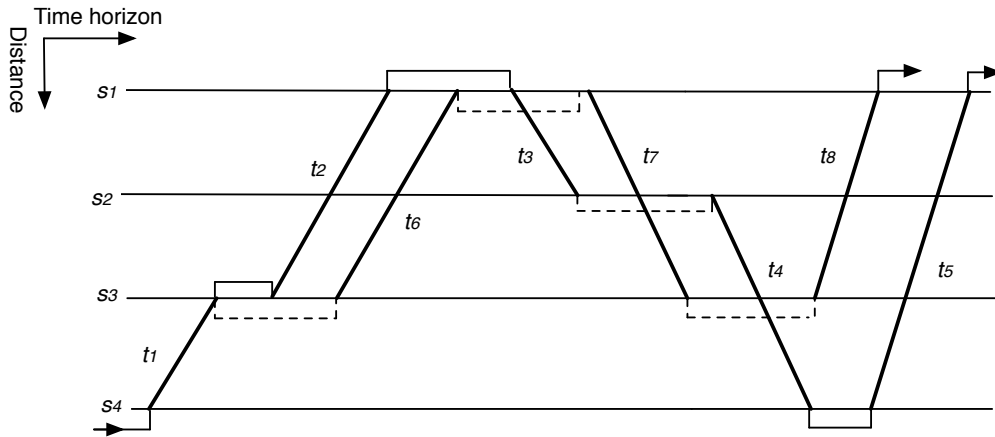


Figure 5.2: Flexible use of train-set

Figure 5.2 show the flexible use of train-set. After the operation of t_1 , the train-set can carry t_2 or t_6 , rather than limited to the section (s_3, s_4) . The train-set circulation can be either $t_1 \rightarrow t_2 \rightarrow t_3 \rightarrow t_4 \rightarrow t_5$ or $t_1 \rightarrow t_6 \rightarrow t_7 \rightarrow t_8$.

The flexible use of train-set increases the utilization of train-set. As long as the requirements of minimum turn around time is met, a train-set has many trains to turn, which will enhance the flexibility of operating a train-set and effectively decrease the number of required train-set. However, it increases the difficulties in scheduling the train-set circulation since the operating sections are quite flexible. In addition, the train-set circulation is more susceptible in case of disturbance, which will lead to more difficulties in rescheduling.

As above analysed, each use of train-set has its own merits and demerits. Both of them are independently or complementary applied in different high-speed lines in China. For example, Beijing-Tianjin HSR and Shanghai-Nanjing HSR lines adopt a complementary use that only some of the train-sets are fixed on certain sections and the others adopt flexible use.

5.2 Terminology and Literature

In this section, we briefly describe the terminologies we use for train-set circulation problem throughout this thesis. The features described are based on the situation at China's HSR. Other operators have similar features, but some features are very characteristic for China's HSR. Having set the terminologies, we give a brief literature overview on train-set routing problem.

5.2.1 Terminology

In this section, we briefly describe the notions of the train-set operations at China's HSR. The concrete models as well as the computational experiments are based on these assumptions.

Train-set and train-set routing problem

Unlike conventional railway, China's HSR mainly uses units instead of locomotive-hauled carriages. Having a driver's seat on both ends of units allows a fast and easy turn-around process in case of direction changes. The physical unit of rolling stock to cover a train is called a *train-set*. Hence, a train-set is a set of passenger cars and power unit(s).

The HSR in China handles train-set operations in a hierarchical manner. A higher level corresponds to the train-sequencing problem of constructing train-set routes from a train timetable. Typically, a train timetable includes millions of potential route sets. A lower level corresponds to the train-set allocation problem of allocating individual train-sets to train routes obtained in the higher level. This hierarchical process simplifies field operations and provides all employees with clarity about their jobs, which is crucial for coordinating operators, especially in a large organization like the China's railway (Chung et al. (2009)). This chapter deals with the train-sequencing problem at the higher level.

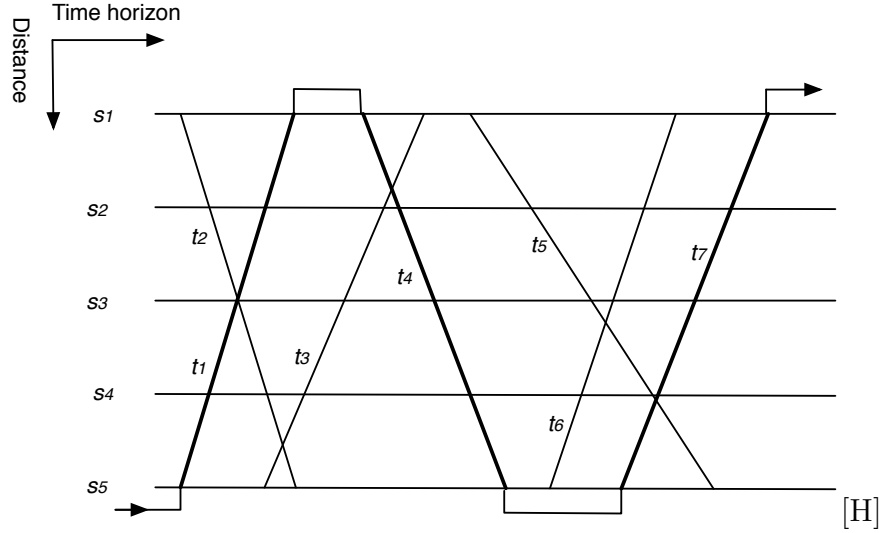


Figure 5.3: A train-set route plan

Moreover, all of the train-sets operated in China are of the highest speed 350 km/h. With the exception sleeping trains, a single train-set type with the same design speed is preferred to be operated in a HSR line (Xie (2010)). Consequently, our problem is simplified to the train-set routing problem without consideration of the train-sets assignment.

A *train-set route* is a sequence of trains with designated departure and arrival times in a train timetable. The *train-set routing problem* is to determine a set of train-set routes covering all trains for a given timetable. For example, t_1 , t_4 and t_7 compose a train-set route in Figure 5.3 and these trains are operated using the same train-set unless any extraordinary failure occurs.

Turn around and turn around time

When a train arrives at a station, it may depart soon as another timetable service, we call this a *turn around*.

Turn around time is the time difference between arrival of the predecessor trains at a terminus, and the departure time of the successor train. If the turn around time is sufficiently long for the train to be cleaned, and for possible shunting operations to be carried out, then both journeys can be operated by the same train-set. In China's HSR, the minimum required time for a turn-around is 15 min. Moreover, the turn around time should not be too long, because long turn around times decrease the efficiency of the train composition utilization.

Splitting and combining of trains

A train-set usually contains different numbers of carriages that can be combined to compositions of various lengths in order to match passenger demand. In China, however, currently operated with only two different train-set types composed of 8 passenger cars

and 16 passenger cars, respectively.

The most peculiar property of the China's railway network is its heavy workload. Some tracks between major cities are so heavily used that often the 3 minutes security headway exist between the trains. In this case, the arrival platform must be freed up as soon as possible, allowing another train to arrive. Heavy traffic often restricts the possibilities to adjust the compositions during the turn-arounds. In fact, it is preferred that no composition change (including splitting and combining) happens at all.

In addition, in practical operation process, the train-set type of 8 passenger cars is seldom considered and used. Therefore we omit the splitting and combining of trains from the problem and model formulations (Xie (2010)).

5.2.2 Literature Overview

Our goal in this chapter is to integrate train-set circulation into the ATP problem and aims to cover the new timetable with a minimal level train-sets as will be specified in detail later.

Rolling stock circulation problem

In Hong et al. (2009), this problem is summarized as the *rolling stock rostering problem*, i.e. Andereg et al. (2002) or *vehicle scheduling problem*, i.e. Bunte and Kliwer (2010) which, in its general sense, consists of the *capacity allocation problem*, i.e. Ben-Khedher et al. (1998) or the *train length problem*, i.e Andereg et al. (2002) that decides the capacity of trains in the schedule, and the *train-set routing*, i.e. (Ben-Khedher et al., 1998; Pisinger and Ropke, 2007) or the *train assignment*, i.e. Andereg et al. (2002) that assigns the rolling stock to the scheduled trains.

Hong et al. (2009) also introduce that the rostering problem dealing with heterogeneous rolling stocks has been considered in the various studies that can be conveniently categorized as the *locomotive assignment problem*, i.e. Rouillon et al. (2006); Vaidyanathan et al. (2008), the *car assignment problem*, i.e. Lingaya et al. (2002), or the *simultaneous locomotive and car assignment problem*, i.e. Cordeau et al. (2001). The mathematical formulation is typically based on the integer multi-commodity flow problem whose commodities correspond to the rolling stock types that can be assigned to each train (Ziarati et al., 1997; Cordeau et al., 2001).

Some other literatures, such as Nielsen (2011), Maróti (2006), Fiole et al. (2006), Budai et al. (2009), Alfieri et al. (2002) and Peeters and Kroon (2008), also devote to the subject on the rolling stock circulation of train unit, which usually consist of operations that uncoupling units from or coupling units to trains due to efficiency reasons.

As described in Section 5.2.1, the China HSR is, however, currently operated with homogeneous rolling stock. All of the train-sets have the design speed of 350 km/h. Moreover, due to the requirement of organization, combining and splitting of trains is

rare in China HSR. That is, each trip specify a type of train-set beforehand and is performed in such a way that the assignment of train-set to trains can not be changed during a trip. It simplifies our problem into the train-set routing problem, see Andereg et al. (2002). A train-set routing is determined by the number of train-sets and the route, namely the sequence of trains, that each train-set traverses in a time horizon, such as (Hong et al. (2009)) present a two-phased train-set routing algorithm to cover a weekly train timetable with minimal working days of a minimal number of train-sets.

Simultaneous rolling stock planning and timetable scheduling/rescheduling problem

Timetable scheduling, rescheduling and rolling stock routing are three complex optimization problem respectively. As summarized in (Geraets et al. (2004)), during the last years a trend towards the integration of several planning steps has emerged, such as a combination of timetabling and rolling stock planning, or a combination of rescheduling and rolling stock circulation.

Some papers have already done many researches on these combined models, such as D'Ariano et al. (2008b), Flier et al. (2008) and Nielsen et al. (2012) integrate rolling stock circulation into rescheduling problem, and Leon (2003) and Cadarso and Marín (2011) integrate rolling stock planning problem into timetable scheduling problem.

D'Ariano et al. (2008b) consider the problem of managing disturbance in real time. The rolling stock circulation are assumed to be fixed as many other literatures. They model the constraints of rolling stock and passenger connections in a same way.

Flier et al. (2008) relax the assumption of fixed rolling stock circulation, and allow changes in the vehicle schedules if they lead to a better disposition timetable. A model for delay management in which the decisions dealing with the circulations are integrated. This problem is proved to be NP-hard even in very simple networks. A polynomial case is identified and different properties and approaches are suggested in this paper.

Leon (2003) integrate rolling stock planning into cyclic timetable scheduling problem. Starting with the most basic case, in which the rolling stock circulation is fixed beforehand, a model of more freedom in choosing the rolling stock circulation is formulated. However, for the above approach to be applicable, we do need to decide beforehand which pairs of trains are allowed to turn on one another.

5.3 Special Features of Train-set Planning in the Adding Train Paths Problem

Traditionally, the overall railway planning problem may be summarized by the following steps: (i) the railway network design problem, (ii) the line planning problem, the timetabling problem, (iii) the rolling stock assignment, (ix) the rolling stock problem,

and (x) the crew scheduling problem. These five subtasks have been usually solved sequentially, and the train-set circulation are usually determined in the tactical planning phase *after* all of the train lines and timetable have been fixed (Geraets et al. (2004), Cadarso and Marín (2011)), as shown in Figure 5.4.

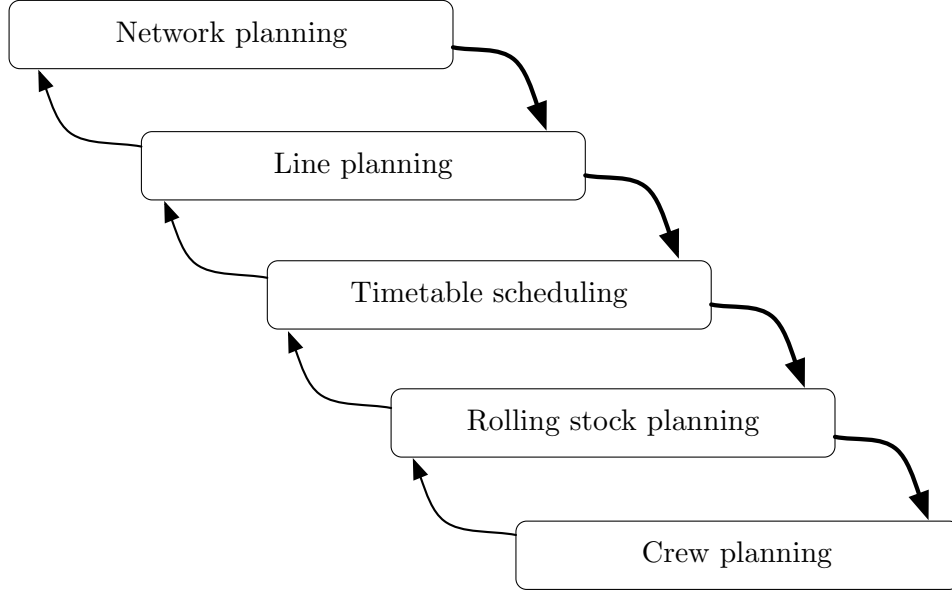


Figure 5.4: Planning phases beforehand

Train-set is a very limited resource when the frequency of train services is very high, especially for the ATP problem. Train-sets has to be scheduled to serve the timetable with ever growing demand for capacity, and the railway company must provide the trains with the adequate train-sets.

However, in the ATP problem, the additional trains do not exist in the initial timetable, and even the number of additional trains depends on the number of instantly available train-sets at the right place. It means, the train-set flow, which imposed by the timetable trips including scheduled and additional trains, is probably not feasible given the limited available of train-sets.

In addition, the schedules of additional trains directly affect the efficiency of train-set. Different allocations will lead to different train-set circulation and appropriate adjustments may reduce the number of required train-sets. As an example, considering the timetable in Figure 5.5, there are 8 trains denoted by t_1 to t_8 . Since the time interval between the arrival of t_4 and the departure of t_6 at station s_3 is less than the minimum turn around time, at least 3 train-sets are required to provide a service of these 8 trains in the considering area.

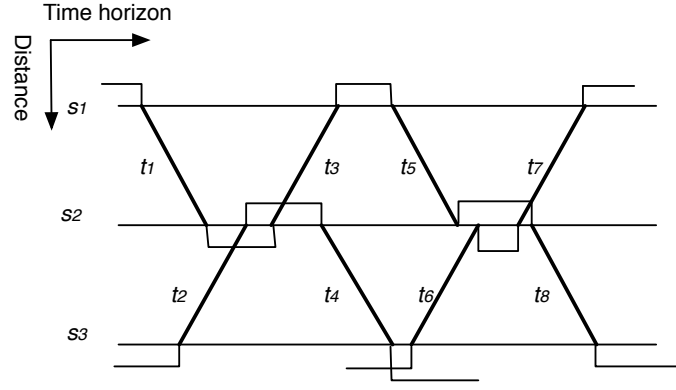


Figure 5.5: 3 required train-sets in the original timetable

However, if t_6 and t_8 are right shifted to meet the minimum turn around time, only two train-sets are required. The corresponding circulations c_1 and c_2 are presented in Figure 5.6.

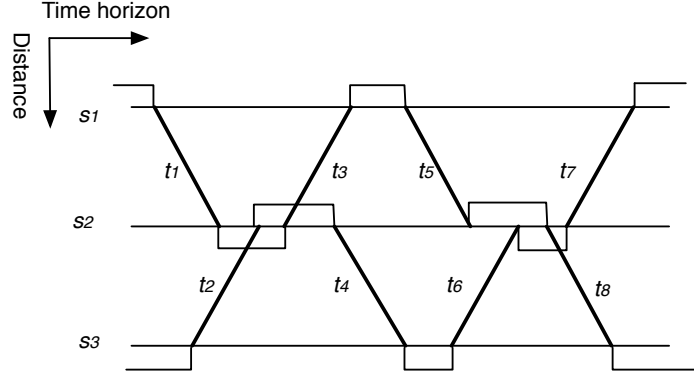


Figure 5.6: 2 required train-sets after adjustment

Based on the above analysis, it is obvious that the schedules of trains have a direct influence towards the use of train-sets. Especially for the ATP problem in China's HSR, the additional trains usually compose of the trains which of low density, and usually depart or arrive concentrately on rush hours. Consequently, the limited available train-set is a crucial constraint for the ATP problem. This has led to an extensive combinatorial research on optimizing extra train paths scheduling and the utilization of the train-sets for additional trains.

In this chapter, our main contribution is that we integrate these planning phases in a model that decides simultaneously additional trains' schedules and train-set circulation for the new timetable.

As analyzed in Chapter 1, the ATP problem may occur both in the phase of tactical and short-term planning. Disturbances for scheduled train-set circulation in these planning phases can not avoid after the existing timetable has been adjusted, and may cause large disruptions for scheduled services when the current circulation of the train-set is changed.

In addition, in the phase of tactical planning to insert non-cyclic trains to schedule a generic timetable, changes to the initial cyclical utilization of train-set might decrease the efficiency and accessibility for maintenance of train-sets. In the phase of short-term planning to insert additional trains for adaption of the increased traffic demand, changes to the scheduled train-set circulation may require coordination with the local traffic controllers of the infrastructure managers to ensure that proposed shunting operations and other local issues are possible (Maróti (2006)).

Hence, for the sake of avoiding large disruptions to the scheduled services and solving adding paths problem within a reasonable computational force, the train-set planning in the ATP problem is decomposed into two sub-problem, as shown in Figure 5.7,

- (1) for the initial timetable, the current train-set circulation is assumed to be fixed beforehand. It is solved as an rescheduling problem with a tight constraint, that a train-set operates the same existing trains in the same sequence as it is scheduled in the initial timetable.
- (2) for the additional trains, the train-set circulation problem is equivalent to covering all the additional trains with minimal number of train-sets.

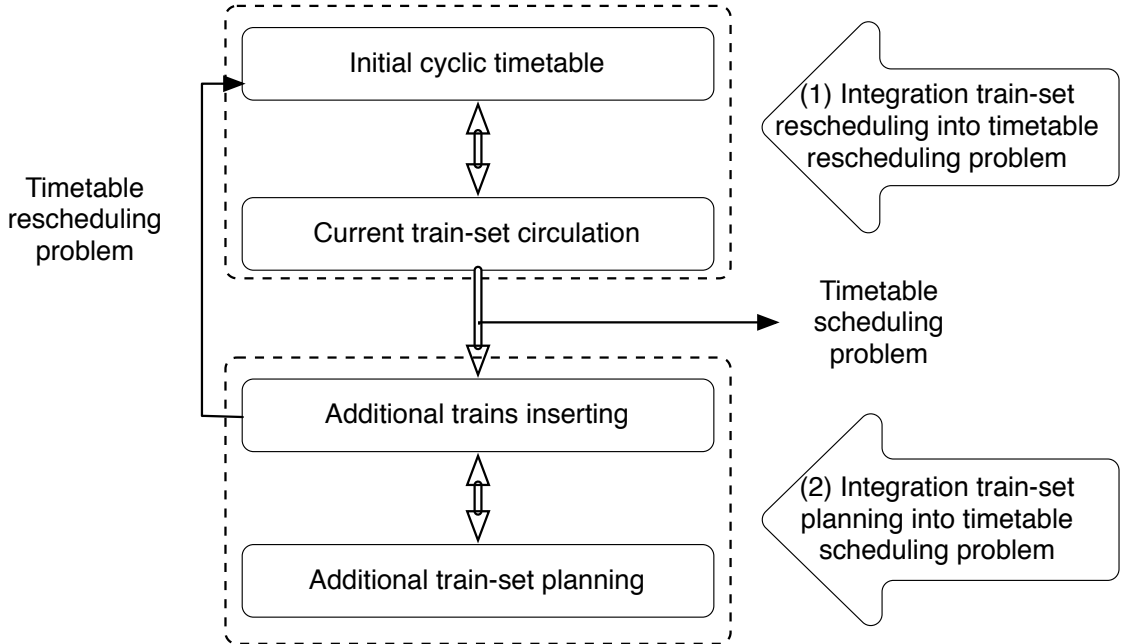


Figure 5.7: Planning phases covered by adding paths problem with our contribution

Then, the integrated ATP problem and train-set circulation can be stated as follows:
given the expected numbers of extra trains, and accounting for the limited available train-sets, find the optimal insertion based on the existing timetable to minimise the adjustments for initial schedules and minimize the required train-sets for the inserted extra trains.

5.4 Formulation for Train-set Routing in Adding Paths Problem

In order to solve the train-set circulation in the ATP problem, let us denote by

$$\begin{aligned} V_{start} &:= \{i \in D : i \text{ is the first event of a train trip}\} \\ V_{end} &:= \{i \in A : i \text{ is the last event of a train trip}\} \end{aligned}$$

where a *train trip* or trip means a particular service specified by origin and destination stations along with intermediate stations, the departure and arrival times of each station, and the type of a train-set assigned to each trip.

All potential combinations between two trips now can be defined and included in the event-activity network. We hence define a new type of activity:

- *Circulation activities* E_{circ} is a set of arcs of the form (u, v) for each $u \in V_{end}$ and $v \in V_{start}$, and models the turn around time which two trips can be operated consecutively within the same circulation.

Note that a circulation activity between event u and v is only possible if the type of the two train-set required for u and v is compatible, and the time interval between u and v is not less than the minimal turn around time L_e .

For sub-problem (1), we can add constraints to keep the current scheduled train-set circulation to the initial timetable:

$$x_j - x_i \geq L_e \quad \forall e = (i, j) \in E_{circ}^{ini} \quad (5.1)$$

For sub-problem (2), we introduce an concept of *rotation* to solve the train-set circulation problem for additional trains.

The term *rotation* is widely used in the airline industry (Lloyd et al. (1997)). The aircraft rotation problem is to determine the routes flown by each aircraft in a given fleet. It is also can be adopted as well in railway system. By a rotation, we mean a circulation in which, as the time horizon repeats, a single train-set covers all the trains. This means every train-set, in the long run, covers the same set of routes. A rotation is a desirable practice in that it maintains train-sets and rails in a homogeneous condition.

To explain a rotation, we consider the example in Figure 5.8. This simple train network shows 4 trains assigned to a line group in which the trains can be operated by the same train-set. The train-set rotation problem orders these trains. For the trains in Figure (5.8), one possible rotation is to cover train t_1 , t_2 , t_3 and t_4 , then repeat the sequence. We represent this simply by $t_1 - t_2 \rightsquigarrow t_3 - t_4 \rightsquigarrow t_1$. The other possible rotation is $t_1 - t_2 \rightsquigarrow t_1 \circ t_3 - t_4 \rightsquigarrow t_3$. Each of these rotations has different characteristics which will be explained now along with the symbols that describe the rotations.

The schedules in this paper are daily. The train trips that are shown in Figure 5.8 are repeated on subsequent days. For rotation $t_1 - t_2 \rightsquigarrow t_3 - t_4 \rightsquigarrow t_1$, on day 1 train-set c_1

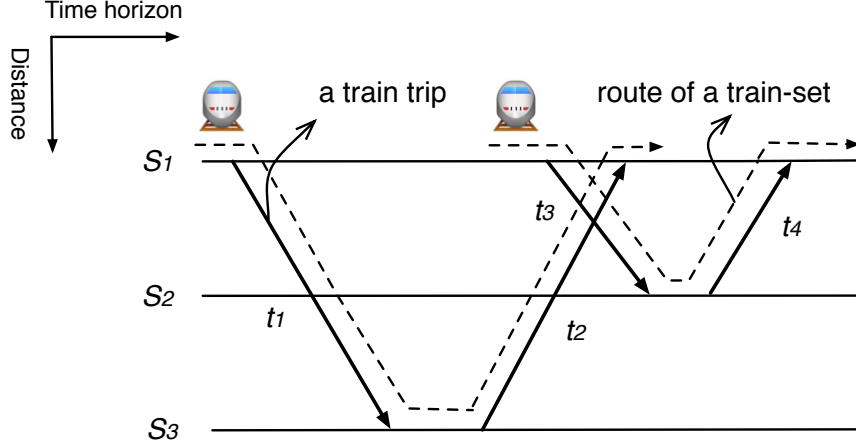


Figure 5.8: A simple train network

takes t_1, t_2 and then spends the night in station s_1 . On day 2, train-set c_1 takes train trips t_3 and t_4 and train-set c_2 takes trains t_1 and t_2 . This complete the rotation. This rotation covers the 4 train trips with two train-sets. The symbol \rightsquigarrow indicates an overnight stay between trains, and consequently the number of overnights is the number of train-sets required to complete the rotation.

The rotation $t_1 - t_2 \rightsquigarrow t_1 \circ t_3 - t_4 \rightsquigarrow t_3$ indicates a different sequence. Train-set c_1 covers trains t_1 and t_2 , overnights in s_1 and then repeats train trips t_1 and t_2 in the next day. Train-set c_2 covers trains t_3 and t_4 , overnights in s_1 too and repeats the same trains in the next day. This rotation uses the same number of train-sets as the previous rotation. What causes the rotations to differ is the choice made overnight at s_1 where they are 2 train-sets. When the train trips covered by one train-set are separate from the trips covered by another train-set, we follow the airline terminology of calling this a *broken rotation* (Lloyd et al. (1997)). The symbol \circ indicates where coverage breaks into different sets of train-set. We refer to these breaks as “continuity breaks” since they break the continuous flow of an aircraft around a cycle that contains all of the train trips. There can be several continuity breaks in a rotation and within each break there can be many overnights. This paper follows the practice of an airline that forbids continuity breaks (Lloyd et al. (1997)).

Using the definition of rotation, we can now define feasible and optimal train-set rotation for additional trains. The train-set rotation problem is to cover all of the new trains with minimal required train-sets.

To assign train-sets to additional trains, we introduce 2 binary decision variables as following,

$$q_{ij} = \begin{cases} 1 & \text{if } (i, j) \in E_{circ} \text{ is chosen,} \\ 0 & \text{otherwise} \end{cases} \quad (5.2)$$

and

$$u_{ij} = \begin{cases} 1 & \text{if } x_j - x_i \geq L_e, \\ 0 & \text{otherwise} \end{cases} \quad (5.3)$$

for all potential circulation activities $e = (i, j) \in E_{circ}$. We also use a binary constant

$$st_{ij} = \begin{cases} 1 & \text{if } s(i) = s(j), \\ 0 & \text{otherwise} \end{cases} \quad (5.4)$$

to denote whether event i and j take place at the same station or not, for $\forall i \in V_{end}$ and $\forall j \in V_{start}$.

In order to obtain feasible rotation for additional trains, we require that for each $i \in V_{end}^{add}$ ($j \in V_{start}^{add}$), exactly one circulation activity starting in i (ending in j) has to be respected. Then the follows constraints are introduced:

$$\sum_{i \in V_{end}^{add}} q_{ij} = 1 \quad \forall i \in V_{end}^{add}, \forall j \in V_{start}^{add} \quad (5.5)$$

$$\sum_{j \in V_{start}^{add}} q_{ij} = 1 \quad \forall i \in V_{end}^{add}, \forall j \in V_{start}^{add} \quad (5.6)$$

Let e_{ij} denote the length of arc $e = (i, j) \in E_{circ}$ which indicates the actual turn around time from i to j , then

$$x_j - x_i \geq L_e - (1 - u_{ij})M \quad (5.7)$$

$$e_{ij} \geq M(1 - st_{ij}) + T_{hor}(1 - u_{ij}) + (x_j - x_i) \quad (5.8)$$

$$e_{ij} \geq L_e \quad (5.9)$$

Constraints (5.7 - 5.8) state that

$$e_{ij} \geq \begin{cases} x_j - x_i & \text{if } s(i) = s(j), x_j - x_i \geq L_e \\ T_{hor} + x_j - x_i & \text{if } s(i) = s(j), x_j - x_i < L_e \\ M & \text{if } s(i) \neq s(j) \end{cases}$$

which means,

- (1) if the time difference between arrival event i and departure event j respects to the minimum required turn around time of train-set, i.e. $x_j - x_i \geq L_e$, then the actual turn around time from i to j holds $e_{ij} \geq x_j - x_i$.
- (2) if the minimum turn around time is not satisfied, i.e. $x_j - x_i < L_e$, then after the end of i , the train-set has to stop at station $s(i)$ for a time horizon (such as an overnight stop) and then it can turn to j . Consequently, the actual turn around time of train-set in this circumstance respects $e_{ij} \geq T_{hor} + x_j - x_i$. In China's HSR, each trip usually has a closed circulation of train-set on a daily basis, therefore the time horizon $T_{hor} = 1440(\text{min})$ in this paper.

- (3) If i and j do not occur at the same station, i.e. $s(i) \neq s(j)$, to connect i and j a train-set goes from one station to another as an empty train. Empty trains are to be avoided as much as possible, and note that empty trains occur in tactical rolling stock circulations quite rarely. That is, we run the model without empty trains by setting $e_{ij} = M$ when $s(i) \neq s(j)$, where M is a large enough constant as defined previous.

Constraints (5.9) specify that the minimal turn around time must be respected for each circulation activity.

By a rotation, a train-set covers all the trains as the time horizon repeats. The circulation activities E_{circ} combined with train activities $E_{train} = E_{trip} \cup E_{dwell}$ corresponds to a train cycle C in Figure 5.9.

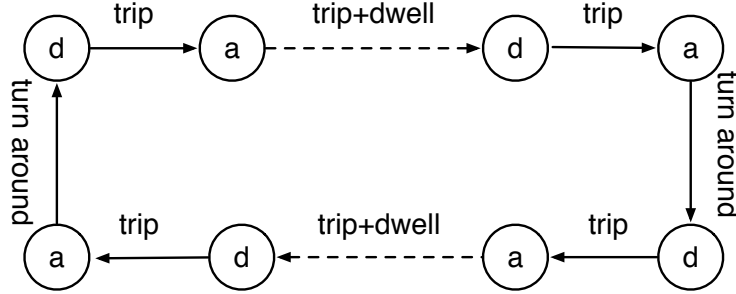


Figure 5.9: A train cycle

Because C is a directed cycle for each set of specified $(i, j) \in E_{circ}$, then

$$TC = \sum_{(i,j) \in E_{circ}} e_{ij}q_{ij} + \sum_{(i,j) \in E_{train}} (x_j - x_i)$$

can be seen as the time it takes a train-set to “travel along the cycle C ” (Leon (2003)). The number of required train-sets to finish the train cycle in a time horizon then can be calculated by $N_{ts} = TC/T_{hor}$, more specifically,

$$N_{ts} = \left[\sum_{(i,j) \in E_{circ}} e_{ij}q_{ij} + \sum_{(i,j) \in E_{train}} (x_j - x_i) \right] / T_{hor} \quad (5.10)$$

One of our objective is to minimize N_{ts} . We can now state the train-set rotation problem in the ATP problem as,

$$\text{(TSR)} \quad \min \quad N_{ts} \quad (5.11)$$

$$\begin{aligned} \text{subject to} \quad (5.1) \quad & \forall i \in V_{end}^{ini}, \forall j \in V_{start}^{ini} \\ (5.5) - (5.9) \quad & \forall i \in V_{end}^{add}, \forall j \in V_{start}^{add} \\ q_{ij}, u_{ij} \in \{0, 1\} \quad & \forall i \in V_{end}^{add}, \forall j \in V_{start}^{add} \end{aligned}$$

The classical train-set routing problem is to determine the specific route flown by each train-set based on a given schedule, that the timetable specifies the departure and arrival times of the trips as well as the actual turn around time of each potential circulation activity. However, in adding paths problem, the additional trains do not exist, and even the number of additional trains depends on the number of instantly available train-sets at the right place. It means in the function (5.10), besides the route of train-set q_{ij} , the departure time x_j , the arrival time x_i and the actual turn around time e_{ij} are all decision variables. Consequently, (5.10) is an non-linear function, which would make the train-set routing problem computationally.

5.4.1 Fixed Train-set Route

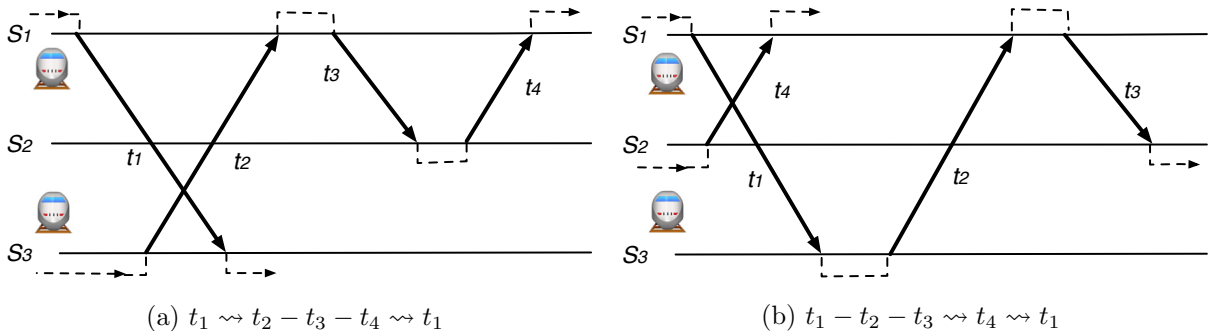
In order to insert additional trains with minimum number of required train-sets and solve the problem within an reasonable time, this section describes a relaxation, and shows how to use it to linearize the train-set objective function in Model (TSR).

We start with a description of the most basic case, in which the route of train-set is fixed beforehand. This fixation is such that the route of train-set, namely the sequence of train trips that each train-set traverses is predefined arbitrarily, and then we can find the optimal inserting solutions with minimum number of train-sets based on the fixed train-set route pattern. See the example in Figure 5.8, if train-set route $t_1 \rightarrow t_2 \rightarrow t_3 \rightarrow t_4 \rightarrow t_1$ is chosen, we denote these *fixed turn around activities* as E_{circ}^{fix} . Then

$$\{(t_1, t_2), (t_2, t_3), (t_3, t_4), (t_4, t_1)\} \subseteq E_{circ}^{fix}$$

is preselected to indicate the route of train-set and their corresponding q_{ij} are set to be 1, else $q_{ij} = 0$.

We emphasize that when additional trains could be scheduled randomly in a time horizon, an arbitrary train-set route would not have an effect on insertion pattern and the train-set rotation. That is, in the ATP problem, it turns out that it is usually not a problem if a “wrong” train-set route is fixed, there seems to be a lot of flexibility in finding insertion solutions for additional trains and the corresponding rotation to minimize the number of train-sets. For example, Figure 5.10 illustrates various insertion solutions and train-set rotations according to the same route of $t_1 \rightarrow t_2 \rightarrow t_3 \rightarrow t_4 \rightarrow t_1$. All of these rotations require at least two train-sets to cover these 4 trains.



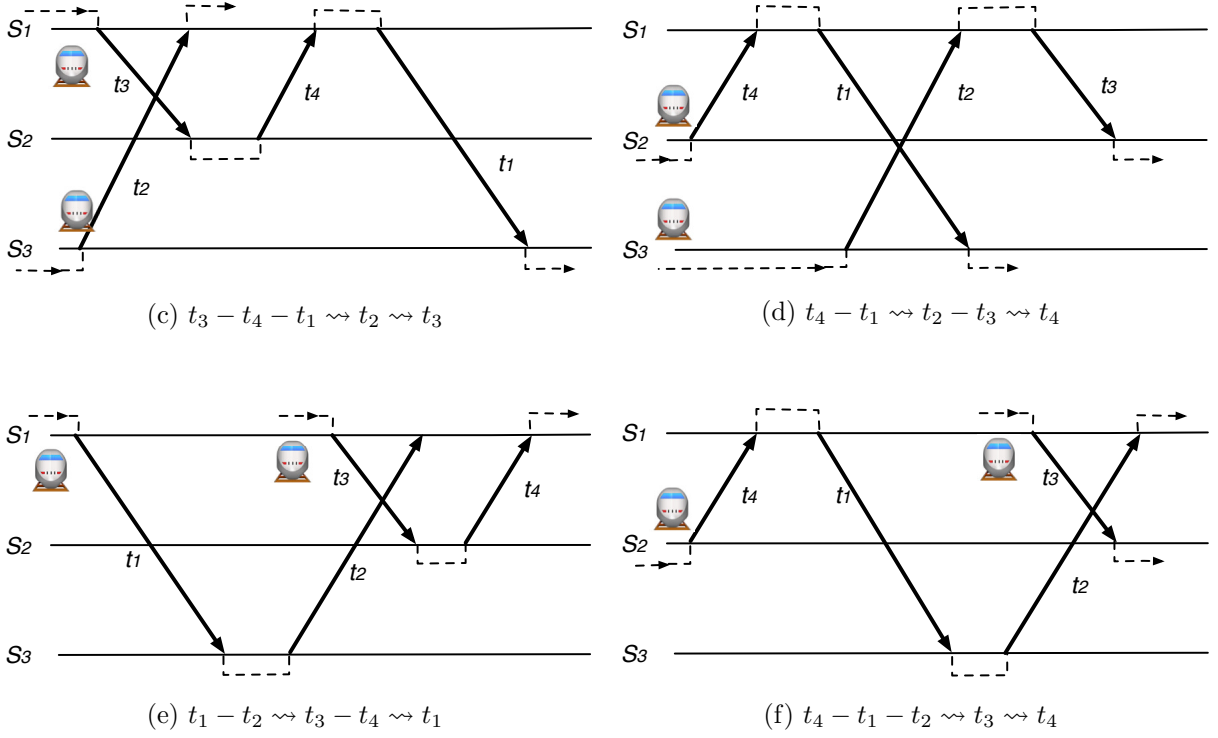


Figure 5.10: Examples of various insertion solutions and train-set rotations to the same route of $t_1 \rightarrow t_2 \rightarrow t_3 \rightarrow t_4 \rightarrow t_1$.

Then by fixing the route, i.e. q is a set of constants, the number of required train-sets N_{ts}^{fix} is

$$N_{ts}^{fix} = \left[\sum_{(i,j) \in E_{circ}} e_{ij} + \sum_{(i,j) \in E_{train}} (x_j - x_i) \right] / T_{hor} \quad (5.12)$$

Consequently, the Model (TSR) can be linearized to

$$\text{(TSR-fixed)} \quad \min \quad N_{ts}^{fix} \quad (5.13)$$

subject to:

$$x_j - x_i \geq L_e \quad \forall (i, j) \in E_{circ}^{ini} \quad (5.14)$$

$$x_j - x_i \geq L_e - (1 - u_{ij})M \quad \forall (i, j) \in E_{circ}^{fix} \quad (5.15)$$

$$e_{ij} \geq T_{hor}(1 - u_{ij}) + (x_j - x_i) \quad \forall (i, j) \in E_{circ}^{fix} \quad (5.16)$$

$$e_{ij} \geq L_e \quad \forall (i, j) \in E_{circ}^{fix} \quad (5.17)$$

$$u_{ij} \in \{0, 1\} \quad \forall (i, j) \in E_{circ}^{fix} \quad (5.18)$$

The function (5.13) simply adds the minimal number of train-sets needed for additional trains with a fixed train-set route. Then Model (TSR-fixed) combined with Model (ATP) described in Chapter 4 are to find a optimal insertion solution with purposes that minimize the deviations to the initial timetable and simultaneously minimize the number of required train-sets.

5.4.2 Flexible Train-set Route

Unlike the simple example presented in Figure 5.10, the insertion of additional trains is influenced by various constraints, such as existing trains in the timetable and reasonable departure-arrival time window for passenger trains, which tends to make the additional trains can not be inserted randomly in the time horizon. This will lead to many overnight stops and result in extra train-sets with a fixed route.

To explain this circumstance, we present the example in Figure 5.11. Here, for the sake of clarity, existing trains and intermediate stations are omitted. 8 additional trains denoted from t_1 to t_8 are needed to be inserted. t_1, t_3, t_5 and t_7 are trips from terminus s_1 to terminus s_2 , and others the reverse journeys from s_2 to s_1 . t_2 and t_3, t_4 and t_5 , and t_6 and t_7 are supposed to start from the corresponding origin station in the time window of $[8 : 00, 9 : 00]$, $[9 : 00, 10 : 00]$, and $[10 : 00, 11 : 00]$ respectively due to the constraints of existing trains and predefined reasonable departure range. If train-set flows the fixed route

$$t_1 \rightarrow t_2 \rightarrow t_3 \rightarrow t_4 \rightarrow t_5 \rightarrow t_6 \rightarrow t_7 \rightarrow t_8 \rightarrow t_1$$

Applying Model (TSR-fixed), we can get the insertion as showed in Figure 5.11, and the corresponding train-set rotation

$$t_1 - t_2 \rightsquigarrow t_3 - t_4 \rightsquigarrow t_5 - t_6 \rightsquigarrow t_7 - t_8 \rightsquigarrow t_1$$

The objective of Model (TSR-fixed) indicates that the route requires $N_{ts}^{fix} = 4$ train-sets to cover all of the additional trains.

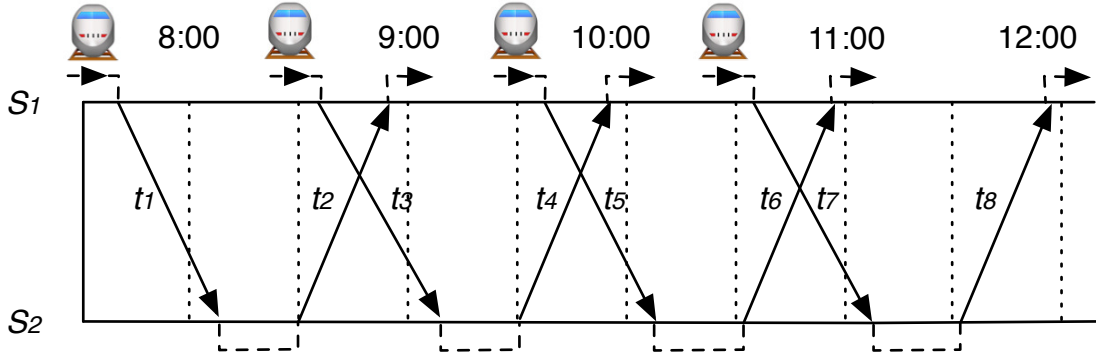


Figure 5.11: A example with fixed train-set route

However, obviously only 2 train-sets can cover the same insertion pattern, if train-set travels the route

$$t_1 \rightarrow t_2 \rightarrow t_5 \rightarrow t_6 \rightarrow t_3 \rightarrow t_4 \rightarrow t_7 \rightarrow t_8 \rightarrow t_1$$

and the corresponding rotation is

$$t_1 - t_2 - t_5 - t_6 \rightsquigarrow t_3 - t_4 - t_7 - t_8 \rightsquigarrow t_1$$

The increase of required train-set in Model (TSR-fixed) resulting from the fixed route omits the potential turn around connections. Practically, the route of a train-set is more likely to cross with the route of another at the same stations within the time intervals that are large enough. Then the sub-routes of arbitrary pairs of train-sets with overnight stop can be interchanged to yield a new routing employing the same level of train-set. This implies that there are many alternative solutions with the same objective value, or even gain a better objective value.

In Figure 5.11, with the predefined route, train-set for $t_2 \rightsquigarrow t_3$, $t_4 \rightsquigarrow t_5$, $t_6 \rightsquigarrow t_7$ and $t_8 \rightsquigarrow t_1$ have overnight stops at s_1 . There are opportunities for arbitrary pairs of train-sets to turn around each other to obtain a new route. Here, t_2 can turn to t_5 , t_7 or t_1 ; t_4 can turn to t_7 , t_1 or t_3 ; t_8 can turn to t_3 , t_5 or t_7 . We denote these *alternative turn around activities* as E_{circ}^{alt} , then

$$\{(t_2, t_5), (t_2, t_7), (t_2, t_1), (t_4, t_7), (t_4, t_1), (t_4, t_3), (t_8, t_3), (t_8, t_5), (t_8, t_7)\} \subseteq E_{circ}^{alt}$$

Moreover, it should be noted that to the specific insertion pattern in Figure 5.11,

- (1) the selection of $(t_2, t_7) \in E_{circ}^{alt}$ will decrease the number of train-sets by 1. Total of $N_{ts} = 4 - 1 = 3$ train-sets will be required for the new route,

$$t_1 \rightarrow t_2 \rightarrow t_7 \rightarrow t_8 \rightarrow t_3 \rightarrow t_4 \rightarrow t_5 \rightarrow t_6 \rightarrow t_1$$

and the corresponding rotation is represented as follows and shown in Figure 5.12,

$$t_1 - t_2 - t_7 - t_8 \rightsquigarrow t_3 - t_4 \rightsquigarrow t_5 - t_6 \rightsquigarrow t_1$$

- (2) the selection of $(t_2, t_5), (t_4, t_7) \in E_{circ}^{alt}$ will decrease the number of train-sets by 2. Total of $N_{ts} = 4 - 2 = 2$ train-sets will be required for the new route,

$$t_1 \rightarrow t_2 \rightarrow t_5 \rightarrow t_6 \rightarrow t_3 \rightarrow t_4 \rightarrow t_7 \rightarrow t_8 \rightarrow t_1$$

and the corresponding rotation is represented as follows and shown in Figure 5.13,

$$t_1 - t_2 - t_5 - t_6 \rightsquigarrow t_3 - t_4 - t_7 - t_8 \rightsquigarrow t_1$$

- (3) the other selections in E_{circ}^{alt} are of no help to decrease the number of required train-sets.

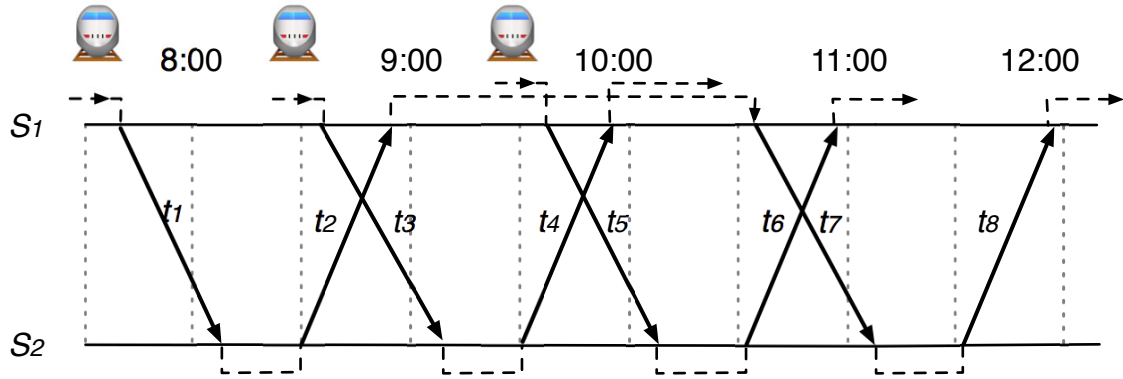


Figure 5.12: The selection of turn around connection (t_2, t_7)

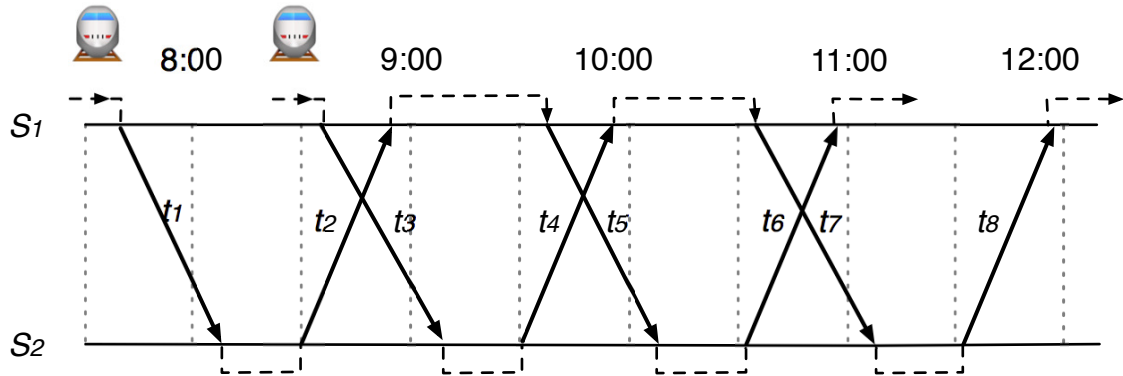


Figure 5.13: The selection of turn around connections (t_2, t_5) and (t_4, t_7)

The above consequences are resulting from that only when the $e \in E_{circ}^{alt}$ which respects the minimal turn around time is chosen, the new route will avoid unwanted overnight stops and then improve efficiency of train-set consequently. Therefore, on the example of Figure 5.11, only activities in

$$\{(t_2, t_5), (t_2, t_7), (t_4, t_7)\} \subseteq E_{circ}^{alt*}$$

have the opportunity to decrease the number of required train-sets. Although it is not necessary in a feasible solution, for each chosen $e \in E_{circ}^{alt*}$ will reduce by one train-set in rotation, as shown in Figure 5.14.

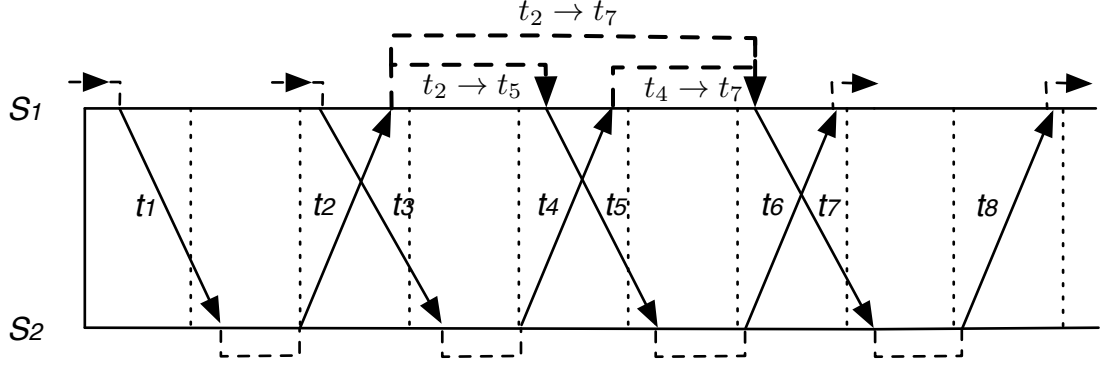


Figure 5.14: Alternative turn around connections on overnight stops

Furthermore, the new train-set route have influences both on the number of train-sets and insertion solution. If the minimum turn around time at s_1 is set to be 1 hour, Figure 5.15 illustrates the inserting solution applying flexible train-set route, in which (t_2, t_5) and (t_4, t_7) are selected, and simultaneously the insertion patter (i.e. the actual time of additional trains) are changed. In this solution, trains t_2, t_4, t_5 and t_7 are adjusted to guarantee the minimum turn around time for t_2 turning to t_5 and t_4 turning to t_7 . As a result, the number of required train-sets decreases by two and then $N_{ts} = 4 - 2 = 2$.

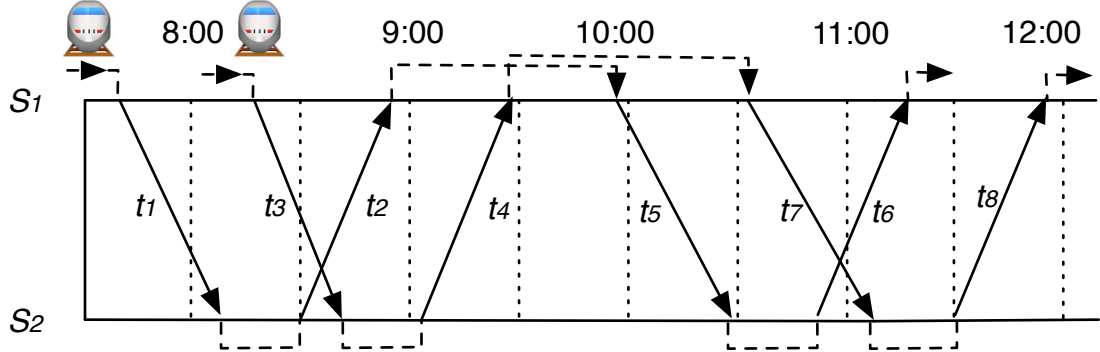


Figure 5.15: Insertion solution based on flexible train-set route

The above procedure is summarized in the following lemmas.

Lemma 5.4.1. Consider a fixed train-set route with predefined E_{circ}^{fix} , any alternative turn around activity $(i, j) \in E_{circ}^{alt}$ can be given by the constraints

$$\begin{aligned} q_{ii'} = 1, u_{ii'} = 0 \quad (i, i') \in E_{circ}^{fix} \\ \& \quad q_{j'j} = 1, u_{j'j} = 0 \quad (j', j) \in E_{circ}^{fix}. \end{aligned}$$

Proof. Note that, (i, i') and (j', j) are arbitrary pairs of turn around activities in fixed train-set route. By assumption E_{circ}^{fix} and constraints (5.2 - 5.3), $q_{ii'} = 1$ illustrates that i

turns to i' in preselected train-set route. $u_{ii'} = 0$ indicates $x_{i'} - x_i < L_e$, that means the time intervals between i and i' is smaller than minimal turn around time, and then $i \rightsquigarrow i'$ has an overnight stop. Similarly, $j' \rightsquigarrow j$ is an overnight stop too. Then $(i, j) \in E_{circ}^{alt}$ is a alternative turn around activity between a arbitrary pair of overnight stops. \square

Definition 5.4.2. Consider an alternative turn around activity $e = (i, j) \in E_{circ}^{alt}$, the set of alternative turn around activities without overnight stops E_{circ}^{alt*} can be given by the constraints

$$u_{ij} = 1 \quad \forall (i, j) \in E_{circ}^{alt},$$

Remark 5.4.1. Note that, $u_{ij} = 1$ represents that the minimal turn around time is respected for train-set turning from i to j , i.e. $x_j - x_i \geq L_e$, which means $i \rightarrow j$ does not have any overnight stop.

Lemma 5.4.3. Consider a fixed train-set route with predefined E_{circ}^{fix} . If an alternative turn around activity $(i, j) \in E_{circ}^{alt*}$ is chosen, the number of required train-sets will reduce by one.

Proof. Consider a fixed train-set route

$$t_1 \rightarrow t_2 \rightarrow \cdots \rightarrow t_n \rightarrow t_1$$

with rotation

$$t_1 - t_2 - \cdots - t_i \rightsquigarrow t_{i'} - \cdots - t_{j'} \rightsquigarrow t_j - \cdots - t_n \rightsquigarrow t_1$$

where $\{(t_i, t_{i'}), (t_{j'}, t_j)\} \in E_{circ}^{alt}$ and $u_{ij} = 1$, then $(t_i, t_j) \in E_{circ}^{alt*}$. If (t_i, t_j) is chosen in the new route which indicates that t_i turns to t_j rather than $t_{i'}$ without overnight stop, then the new rotation

$$t_1 - t_2 - \cdots - t_i \rightsquigarrow t_{i'} - \cdots - t_{j'} \rightsquigarrow \downarrow t_j - \cdots - t_n \rightsquigarrow t_1$$

is equivalent to

$$t_1 - t_2 - \cdots - t_i - t_j - \cdots - t_n \rightsquigarrow t_1 \circ t_{i'} - \cdots - t_{j'} \rightsquigarrow t_{i'}$$

Two overnight stops $(t_i, t_{i'})$ and $(t_{j'}, t_j)$ are avoided, but one additional set of train-sets is added since the coverage breaks into two different sets of train-sets from a single set. Consequently, the number of required train-sets in the new route reduces by $2 - 1 = 1$. \square

Based on the above analysis, we can now formulate the train-set planning problem with flexible route in adding paths problem. Firstly, we introduce new binary variable

$$l_{ij} = \begin{cases} 1 & \text{if } (i, j) \in E_{circ}^{alt*}, \\ 0 & \text{otherwise} \end{cases} \quad (5.19)$$

to indicate that if $(i, j) \in E_{circ}^{alt*}$, then $l_{ij} = 1$, else $l_{ij} = 0$ where $\forall i \in V_{end}^{add}$ and $\forall j \in V_{start}^{add}$.

Remark 5.4.2. According to lemmas (5.4.1 and 5.4.3) and constraints (5.2 and 5.3), $(i, j) \in E_{circ}^{alt*}$ if and only if constraints as follows

$$\begin{aligned}
q_{ii'} = 1, u_{ii'} = 0 & \quad (i, i') & \in E_{circ}^{fix} \\
q_{j'j} = 1, u_{j'j} = 0 & \quad (j', j) & \in E_{circ}^{fix} \\
u_{ij} = 1 & \quad (i, j) & \in E_{circ}^{alt} \\
q_{ii'}, q_{j'j} \in \{0, 1\} & \quad (i, i'), (j', j) & \in E_{circ}^{fix} \\
u_{ii'}, u_{j'j} \in \{0, 1\} & \quad (i, i'), (j', j) & \in E_{circ}^{fix}
\end{aligned}$$

are satisfied. It is equivalent to that when $q_{ii'} + q_{j'j} - (u_{ii'} + u_{j'j} - u_{ij}) = 3$, then $(i, j) \in E_{circ}^{alt*}$.

Hence, constraints (5.19) can be rewrote by q and u

$$l_{ij} = \begin{cases} 1 & \text{if } q_{ii'} + q_{j'j} - (u_{ii'} + u_{j'j} - u_{ij}) = 3, \\ 0 & \text{otherwise} \end{cases} \quad (5.20)$$

where $\forall i \in V_{end}^{add}$ and $\forall j \in V_{start}^{add}$.

Next, the other binary variable is defined as

$$k_{ij} = \begin{cases} 1 & \text{if } (i, j) \in E_{circ}^{alt*} \text{ is chosen,} \\ 0 & \text{otherwise} \end{cases} \quad (5.21)$$

to present circumstances that if alternative turn around $(i, j) \in E_{circ}^{alt*}$ is chosen in the new flexible route, then $k_{ij} = 1$, else $k_{ij} = 0$.

As analyzed in lemma (5.4.3), the number of required train-sets then can be represented as

$$N_{ts} = N_{ts}^{fix} - \sum_{(i,j) \in E_{circ}^{alt*}} k_{ij} \quad (5.22)$$

By introducing additional constraints as follows,

$$k_{ij} \leq l_{ij} \quad \forall i \in V_{end}^{add}, j \in V_{start}^{add} \quad (5.23)$$

$$\sum_{i \in V_{end}^{add}} k_{ij} \leq 1 \quad \forall j \in V_{start}^{add} \quad (5.24)$$

$$\sum_{j \in V_{start}^{add}} k_{ij} \leq 1 \quad \forall i \in V_{end}^{add} \quad (5.25)$$

$$k_{ij} \leq st_{ij} \quad \forall i \in V_{end}^{add}, j \in V_{start}^{add} \quad (5.26)$$

$$k_{ij}, l_{ij} \in \{0, 1\} \quad \forall i \in V_{end}^{add}, j \in V_{start}^{add} \quad (5.27)$$

Constrains (5.23) imply that only the selection of an alternative turn around activity which holds $(i, j) \in E_{circ}^{alt*}$ can reduce the number of train-sets. That is, when $l_{ij} = 0$ (i.e. $(i, j) \notin E_{circ}^{alt*}$), the choice of any turning around operation from i to j can not avoid

an overnight stop and consequently the number of train-sets N_{ts} can not be decreased. Conversely, when $l_{ij} = 1$, k_{ij} takes the value 0 or 1. If and only if i turning around to j as an alternative solution to the fixed train-set route is selected (i.e. $k_{ij} = 1$), then N_{ts} could be reduced by one, since there is no overnight stop in the sub-alternative-route $i \rightarrow j$.

Constraints (5.24-5.25) are similar to constraints (5.5-5.6) and enforce that for every train there is exactly one train-set connection to turn to and be turned by another train. Constraints (5.26) specify that the turning around is forbidden between two operations which are of different stations.

We can now state the train-sets rotation for the ATP problem as

$$\begin{array}{ll} \text{(TSR-alt)} & \min \quad N_{ts} \\ & \text{subject to:} \quad \text{constraints (5.14 -5.18 and 5.20-5.27).} \end{array}$$

Combined with the Model (ATP) formulated in Chapter 4, we can integrate the train-set circulation to the ATP problem to minimize the adjustments to the initial timetables F_a and simultaneously minimize the required train-sets F_s , where $F_s = N_{ts}$.

5.5 Experimental Studies

This section provides details of comprehensive numerical investigations to identify whether good solutions can be obtained using the methodology and techniques proposed in this paper. The primary aim of the adding paths problem is to solve the problem that

How to operate additional trains most appropriately with minimum number of train-sets and without leading large disruption to initial timetable?

Meanwhile, we also would like to know the affecting factors to this problem. For example,

- (1) *What effect the various level of accepted disruption have?*

Different control parameters for tolerance of disruption, composed of *allowable adjustment* Δ and *periodic structure* θ , are investigated in Part 1.

- (2) *What effect the different objective functions have?*

The adding paths problem, using the multi-objective function (1a) = $w_a F_a + w_s F_s$ that minimizing total adjustments and required train-sets, and single-objective function (1b) = F_s that only minimizing required train-sets, are analysed in Part 2.

- (3) *What effect the introduction of time window constraints have?*

With or without time window as each insertion option are compared in Part 3.

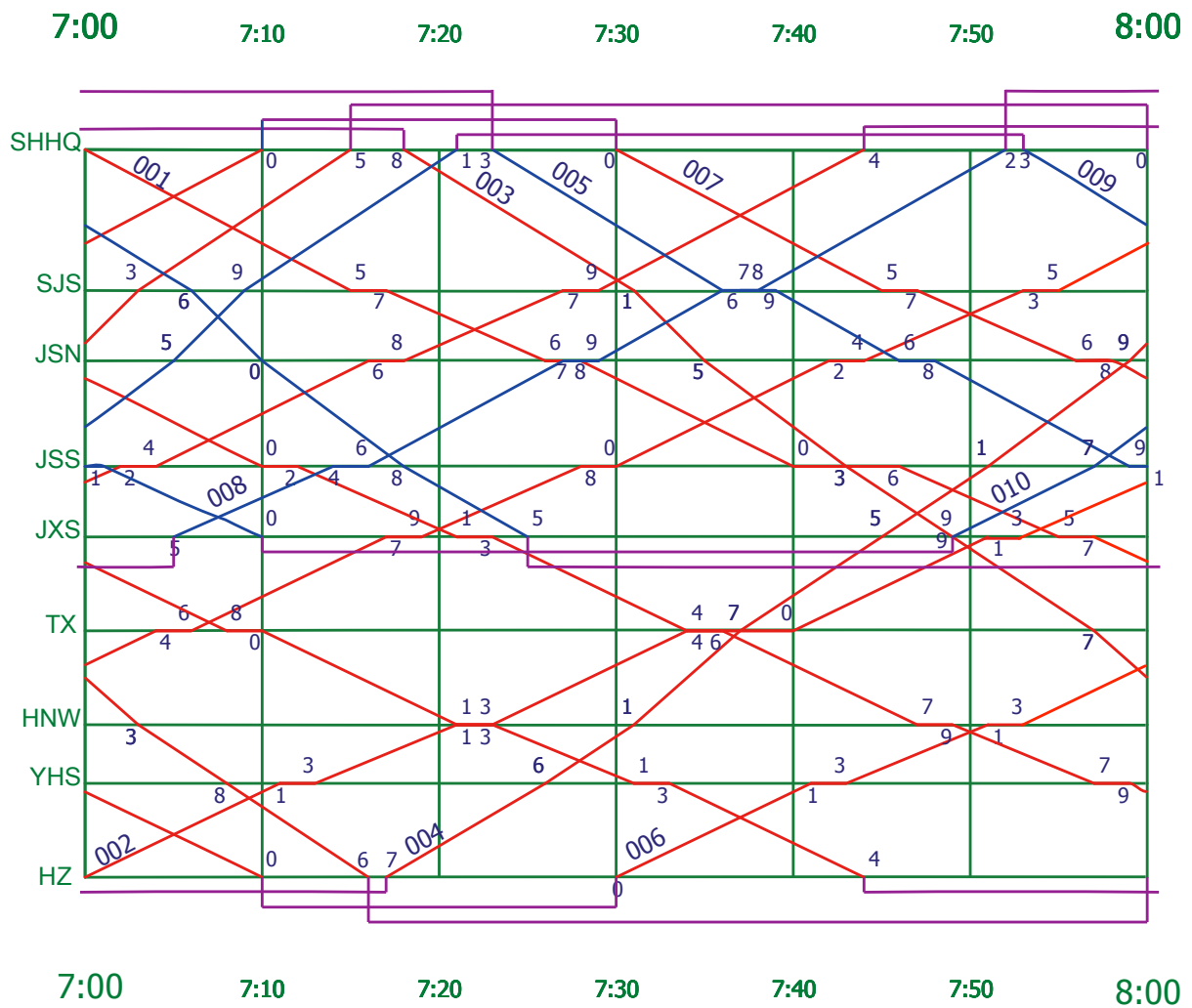
(4) *What effect the different use of train-sets have?*

Two different turn around approaches in the use of train-sets are tested in Part 4.

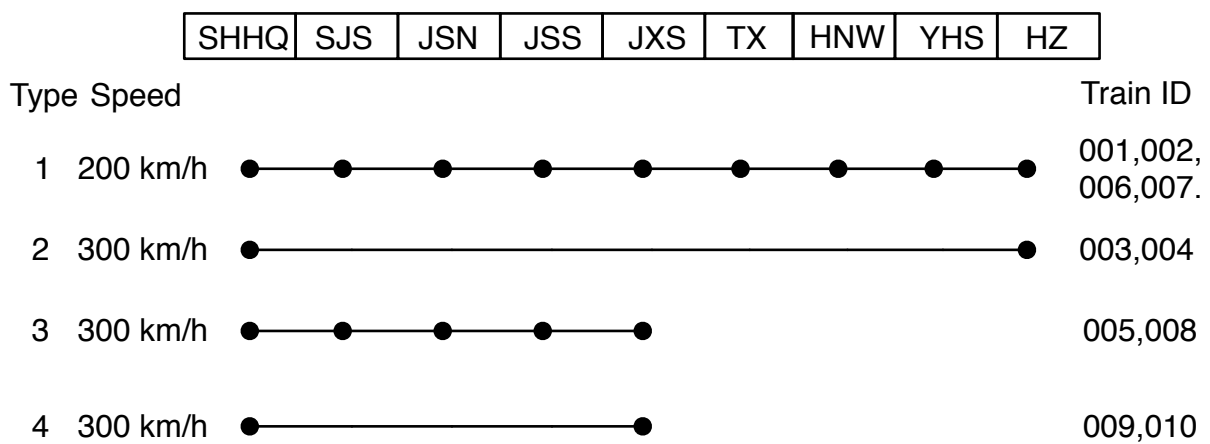
5.5.1 Test Problem

The formulation and the strategies have been applied to Shanghai-Hangzhou high-speed rail line, which consists of double-tracked high-speed railway lines that are the major links connection Shanghai Hongqiao (SHHQ), Songjian South (SJS), Jinshan North (JSN), Jiashan South (JSS), Jiaxing South (JXS), Tongxiang (TX), Haining West (HNW), Yuhang South (YHS) as well as Hangzhou (HZ). In this study, our focus is on a time horizon of a 7 hours generic cyclic timetable in the time period of (6:00 am - 13:00 pm), and it includes 78 passenger trains in both down direction (from SHHQ to HZ) and up direction (from HZ to SHHQ). The cyclic nature of the timetable is illustrated in Figure 5.16a. Four Types of trains are used, see Figure 5.16b:

- type 1: medium-speed trains (200 km/h) which are composed of 2 trains (001,002,006 and 007) in each direction between the railroad region of SHHQ - HZ, and scheduled to stop at every intermediate stations,
- type 2: high-speed trains (300 km/h) which are composed of 1 train (003 and 004) in each direction between the railroad region of SHHQ - HZ, without any scheduled stop at intermediate stations,
- type 3: high-speed trains (300km/h) which are composed of 1 train (005 and 008) in each direction between the railroad region of SHHQ - JXS, and scheduled to stop at every intermediate stations,
- type 4: high-speed trains (300km/h) which are composed of 1 train (009 and 010) in each direction between the railroad region of SHHQ - JXS, without any scheduled stop at intermediate stations.



(a) One hour time-space diagram for the track between Shanghai and Hangzhou



(b) Train types with various trip, speed and stop schedules in initial timetable

Figure 5.16: Time-space diagram for Shanghai-Hangzhou high-speed railway

In the experiments, minimum headways are set to 3 minutes for both consecutive arrivals and departures. Acceleration and deceleration times are set to 2 and 1 min-

utes respectively for both high-speed and medium-speed trains. In addition, taking the variable velocity into consideration, maximum driving time is set to 110% (w.r.t. the minimum driving time). The train-set circulation for existing trains is constructed from the initial timetable.

All the experiments are performed on a PC with Intel(R) Core(TM) i3 CPU 530 @ 2.93GHZ + 2.93GHZ and 8 GB of RAM, and all the algorithms are implemented in Visual Studio 2013 on the Windows 8.1, 64 bit. IBM ILOG Cplex 12.5 with default set is used as a solver. We have run all the instances with a time limit of 1 h, and all the results report the outcome when this time limit was reached (or when a proven optimal solution is found).

5.5.2 Results

Part 1: Experiments with different tolerance of disruption

In this part, two new types of train without time window is planned to insert as extra trains in the initial timetable. It is operated between the railroad region of SHHQ - HZ. The down train is scheduled to stop at intermediate stations SJS, JXS and HNW, while the up train stops at JSN, JXS and YHS at least 2 min, see Figure 5.17. For the sake of simplification, the maximum dwell time is set to 7 min at any arbitrary intermediate station in all experiments. The value of w_s and w_a in objective function (1a) are set to 1000 and 10 respectively. All of the additional trains appear in pairs of reverse directions (i.e. one down train and one up train).

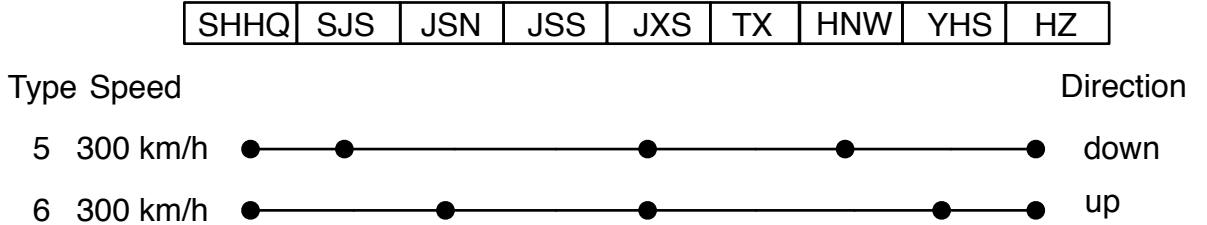


Figure 5.17: Speed and stops schedule for additional trains

The results from the experiments with various tolerance of disruption and a number of 10, 12, 14 and 16 additional trains are tested respectively in Table 5.1. A increase both in the number of additional trains and the level of tolerance generate an increase in computational time predictably.

Besides, the two main aspects considered when analysing the results in Table 5.1 are the differences between the tolerance of the initial timetable and quality of solutions generated. The increasing freedom (i.e. more options to modify the initial timetable) represented by bigger Δ and θ provides as good or better solutions, especially regarding the number of required train-sets for additional trains. The choice of high tolerance has an obvious effect on decreasing the number of train-sets.

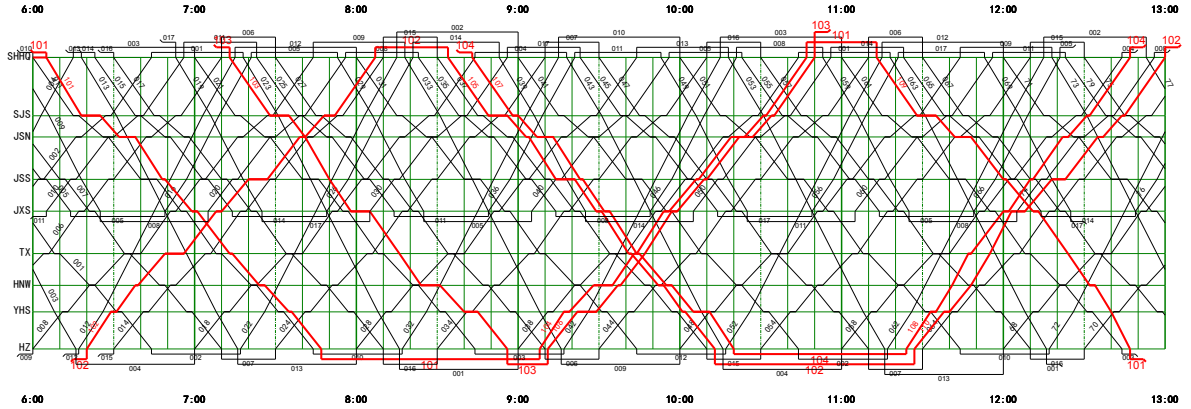
Table 5.1: Results from experiments with different tolerance of disruption

Case	Nr. of \mathcal{T}^{add}	Tolerance		Results			
		Δ	θ	Objective value	Number of train-sets	Adjustments (min)	Time ^a (sec)
1	10	0	0	4000	4	0	10
2		1	0	4000	4	0	41
3		1	1	4000	4	0	49
4		2	0	3100	3	10	1822
5		2	1	3000	2	100	37
6		2	2	3000	2	100	155
7		3	0	3100	3	10	(0.64%)
8		3	1	3000	2	100	194
9		3	2	3000	2	100	172
10		3	3	3000	2	100	195
11	12	0	0	4000	4	0	23
12		1	0	4000	4	0	84
13		1	1	4000	4	0	109
14		2	0	3310	3	31	(5.04%)
15		2	1	3310	3	31	(4.53%)
16		2	2	3310	3	31	(3.92%)
17		3	0	3310	3	31	(6.04%)
18		3	1	3310	3	31	(4.53%)
19		3	2	3310	3	31	(5.74%)
20		3	3	3310	3	31	(6.94%)
21	14	0	0	5000	5	0	96
22		1	0	5000	5	0	188
23		1	1	5000	5	0	256
24		2	0	4200	4	20	(4.76%)
25		2	1	4100	3	110	(2.43%)
26		2	2	4100	3	110	(2.43%)
27		3	0	4210	4	21	(4.98%)
28		3	1	4110	3	111	(12.89%)
29		3	2	4100	3	110	(2.43%)
30		3	3	4100	3	110	(3.82%)
31	16	0	0	6000	6	0	348
32		1	0	6000	6	0	739
33		1	1	6000	6	0	644
34		2	0	4430	4	43	(9.71%)
35		2	1	4520	4	52	(11.50%)
36		2	2	4420	4	42	(9.28%)
37		3	0	4420	4	42	(9.50%)
38		3	1	4440	4	44	(9.91%)
39		3	2	4420	4	42	(9.50%)
40		3	3	4430	4	43	(9.71%)

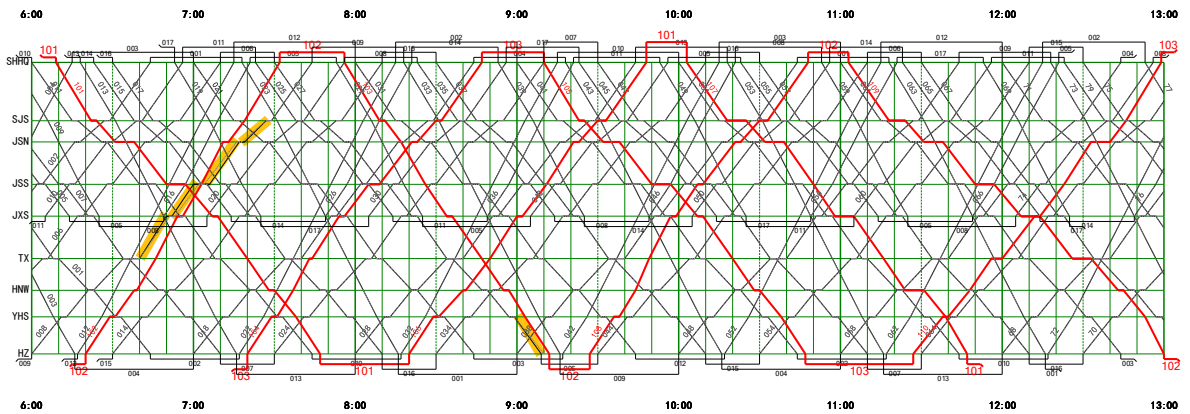
^a The numbers in the parentheses refer to the relative gap when the time limit of 1 h was exceeded without an optimal solution being verified.

Figure 5.18 represents the solutions for inserting 10 trains with different level of tolerance,

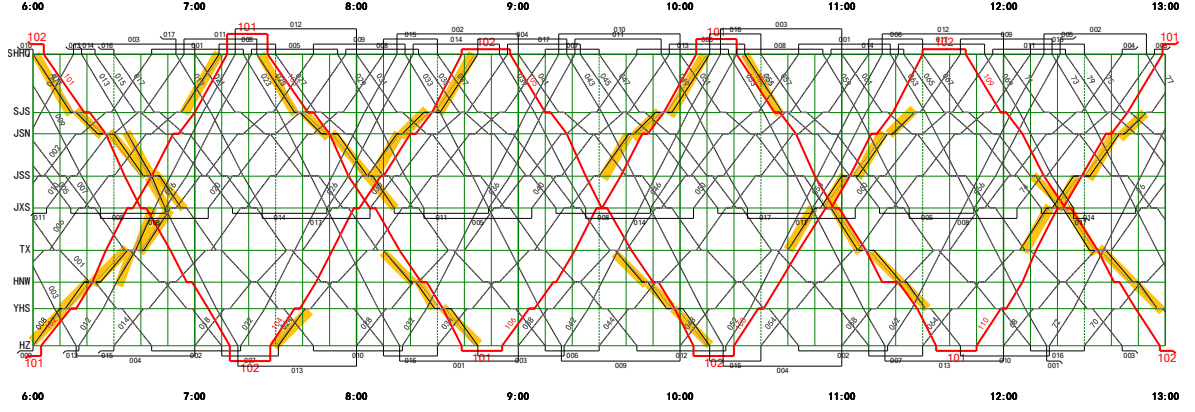
- (1) if the allowable adjustment $\Delta = 0$ and periodic structure $\theta = 0$, which implies the initial timetable is fixed, 4 train-sets are required to cover these 10 additional trains. The insertion and train-set circulation are shown in Figure 5.18a.
- (2) if $\Delta = 2$ and $\theta = 0$, it constraints that the initial trains can be left or right shift at most 2 min but must departure from the corresponding original station at a exact periodic interval. Figure 5.18b demonstrates that 3 train-sets are required when the initial trains are modified 10 min (see the yellow rectangle area). The departures at original station attempt to keep same as initial schedules to prevent large adjustment.
- (3) if $\Delta = 2$ and $\theta = 2$, the initial trains have a higher tolerance both in adjustment and periodic structure. By comparison, only 2 train-sets are sufficient for operating the same trains. The corresponding solution is illustrated in Figure 5.18c. The decrease in train-sets is at the cost of 100 min adjustment to initial timetable (see the yellow rectangle area).



(a) Case 1: 10 additional trains, $\Delta = 0$, $\theta = 0$.



(b) Case 4: 10 additional trains, $\Delta = 2$, $\theta = 0$.



(c) Case 6: 10 additional trains, $\Delta = 2$, $\theta = 2$.

Figure 5.18: Solutions with different level of tolerance

Even though high tolerance decreases the number of train-sets dramatically, on the other hand it impacts on problem size (i.e. the higher tolerance, the more initial trains can be rescheduled and the more options that additional trains can be inserted), and consequently may become more time consuming to solve the problem.

Table 5.1 only specify the best solutions found within a certain time while it also may be relevant to analyse the progress over time. In several of the scenarios presented in Table 5.1, for example when 12 trains are planned to insert, case (14-20) could not verify optimality of the found solutions (where optimal now refers to the optimal solution to the problem formulation of different tolerances) within 1 h, but a solution within an acceptable, relative gap was found. In Table 5.2, the solution progress in case (14, 15, 16 and 19) is presented showing that the same solutions (and corresponding gap) were found within 9 min or less. In comparison with the results in Table 5.1, case (17, 18 and 20) provide similar solutions in 10 min shown in Table 5.2 that require the same train-sets but an extra 2 min, 1 min and 2 min adjustments in case (17), (18) and (19) respectively.

The relative gap in relation to the size of the objective value needs to be considered to provide an appropriate and effective stopping criterion. That is, in the case of $w_s = 1000$ and $w_a = 10$ in objective function (1a), a relative gap of for example 33% may be tolerated for the problem with objective value 3000 that aims to get a solution with minimum train-sets,

Table 5.2: Objective value of the cases 14-20 from Table 5.1 with different time limits

Case	Time limit 3 ^a (min)			Time limit 6 ^a (min)			Time limit 9 ^a (min)		
	Objective value	Number of Train-sets	Relative gap (%)	Objective value	Number of Train-sets	Relative gap (%)	Objective value	Number of Train-sets	Relative gap (%)
14	3310 *	3	9.37	3310	3	9.37	3310	3	8.98
15	3310 *	3	9.37	3310	3	9.37	3310	3	9.37
16	3380	3	11.24	3310 *	3	9.65	3310	3	9.65
17	3820	3	21.47	3410	3	12.02	3330	3	11.76
18	3360	3	10.71	3360	3	10.71	3320	3	9.34
19	4610	4	34.92	3320	3	9.64	3310 *	3	9.37
20	3480	3	13.79	3330	3	9.91	3330	3	9.91

^a The number with “*” refers to that no further improvements were found within 1 h.

Part 2: Experiments using objective function (1b)

In order to also consider the effects of using different objective functions, the same scenarios as for the experiments presented in Table 5.1 have been solved but with objective function (1b) instead of (1a). The results are presented in Table 5.3. The same relation between the tolerance regarding objective value as found using (1a) can be seen here. As might have been expected, same or less train-sets are required when only minimizing the costs are considered. Overall, the computational effort required to solve the problems is decreased dramatically.

Table 5.3: Results from experiments using same scenarios as for experiments presented in Table 5.1, but with objective function (1b)

Case	Total Nr. of \mathcal{T}^{add}	Tolerance		Results	
		Δ	θ	Number of Train-sets	Time (sec)
1	10	0	0	4	10
2		1	0	4	22
3		1	1	4	24
4		2	0	2	28
5		2	1	2	25
6		2	2	2	24
7		3	0	2	36
8		3	1	2	29
9		3	2	2	35
10		3	3	2	24
11	12	0	0	4	51
12		1	0	4	56
13		1	1	4	64
14		2	0	3	25
15		2	1	3	35

Continued on next page

Table 5.3 – continued from previous page

Case	Total Nr. of \mathcal{T}^{add}	Tolerance		Results	
		Δ	θ	Number of Train-sets	Time (sec)
16	14	2	2	3	51
17		3	0	3	45
18		3	1	3	40
19		3	2	3	86
20		3	3	3	38
21		0	0	5	109
22		1	0	5	192
23		1	1	5	218
24		2	0	3	61
25		2	1	3	476
26	16	2	2	3	299
27		3	0	3	167
28		3	1	3	607
29		3	2	3	445
30		3	3	3	456
31		0	0	6	311
32		1	0	6	319
33		1	1	6	425
34		2	0	4	41
35		2	1	4	43
36		2	2	4	116
37		3	0	4	227
38		3	1	4	196
39		3	2	4	287
40		3	3	4	128

Part 3: Experiments of inserting trains with time window constraints

In practice, additional trains are usually supposed to insert in a specific time period, namely time window constraints. For instance, in the application of increasing train services to meet the passenger flow, the additional trains are planned to departure in rush hours. Furthermore, when inserting interline trains, the options of time slot for departures and arrivals are very limited, even fixed at a precise time generally.

For sake of simplification, all of the initial trains are fixed and the considered time horizon is divided into 3 independent time period, i.e. 6:00-8:00, 8:00-10:00 and 10:00-13:00. The number of required new trains in each time period is shown in Table 5.4, and as well a comparison results of inserting trains with and without time window constraints. With time window constraints, the number of train-sets increases as expected due to imbalance utilization. However, the computational time dramatically decreases to less than 0.5 s even inserting 16 trains. The time window constraints not only narrow the search space of insertion, but also cut the option of circulation down.

Table 5.4: A comparison of inserting additional trains with and without time window

Total Nr. of \mathcal{T}^{add}	With time window					Without time window	
	Nr. of \mathcal{T}^{add} in each time window			Nr. of train-sets	Time (sec)	Nr. of train-sets	Time (sec)
	6:00-8:00	8:00-10:00	10:00-13:00				
10	4	4	2	4	0.15	4	10
12	4	6	2	6	0.18	4	23
14	4	6	2	6	0.29	5	96
16	6	8	2	8	0.43	6	348

Part 4: Experiments using different approaches on the use of train-set

The ATP model proposed in this paper is adapted to both fixed and flexible use of train-set, as introduced in Section 5.1. With the purpose of analysing the impacts of different uses of train-sets on the ATP problem, 5 types of train with various railroad sections are planned to insert in this experimental part. Figure 6.1 indicates the information of speed, stop schedule, running direction and railroad sections, and the number as well for each type of train.

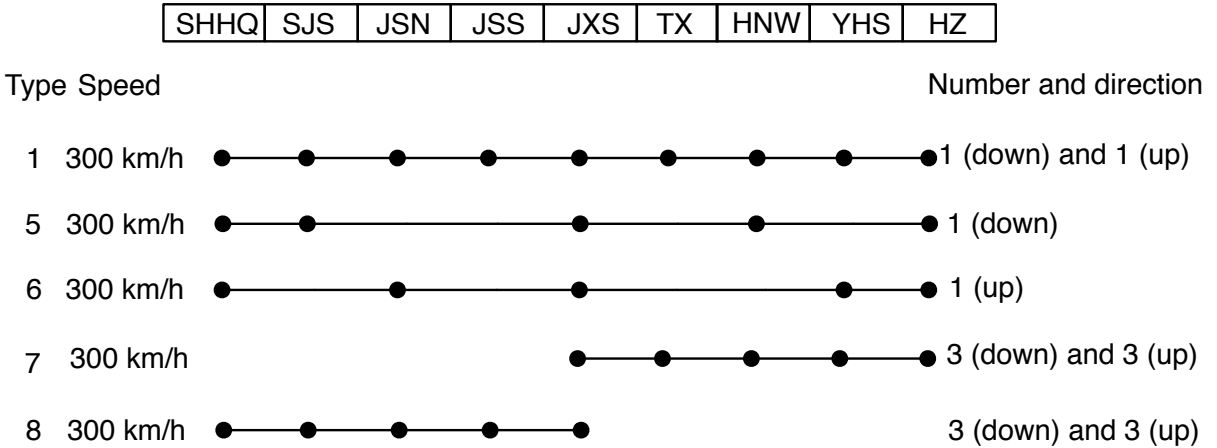


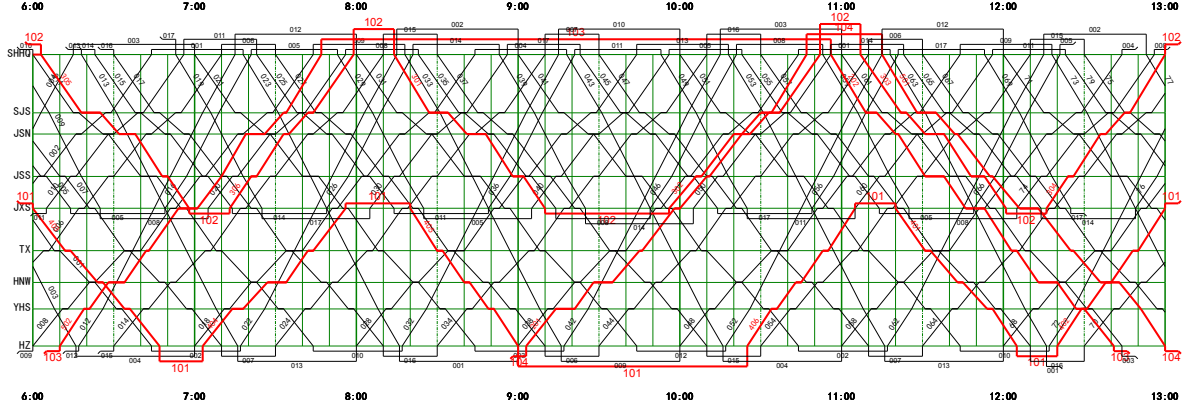
Figure 5.19: Information of additional trains

The solution using different approaches of train-set application is illustrated in Figure 5.20. All of the initial trains are supposed to be fixed. Figure(5.20a) shows that using train-set in certain operating sections, 4 train-sets are required to cover all of the 16 additional trains. Train-set 101 carries train 302, 303, 304, 305, 306 and 301 successively in operating sections SHHQ-JXS; train-set 102 carries 403, 404, 405, 406, 401 and 402 in operating sections JXS-HZ; train-sets 103 and 104 carry in operating sections SHHQ-HZ.

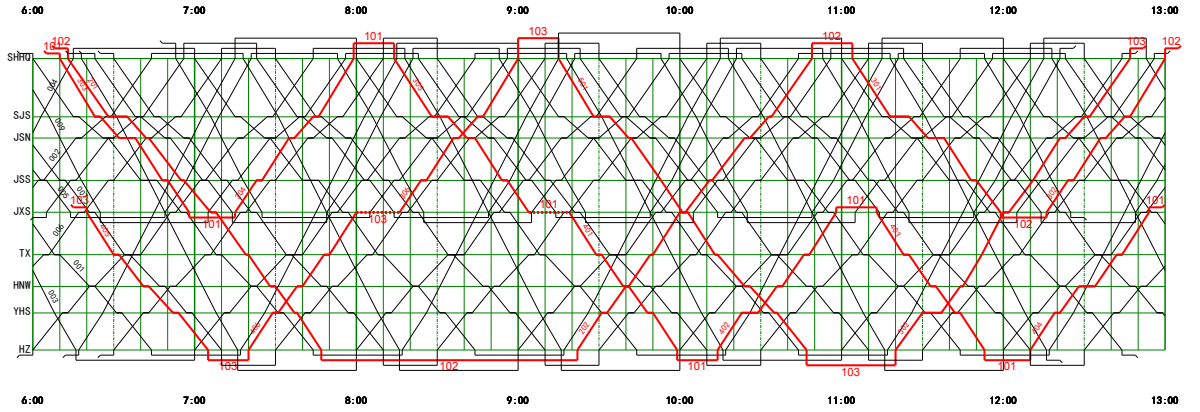
However, when using train-set in uncertain operating sections, see Figure 5.20b, only 3 train-sets are required. In order to be distinguished from dwell activities, the turn around activities between trains of same direction are represented by red dotted line in Figure 5.20b, such as train-set 103 at station JXS where turning from train 406 to 306.

The flexible use of train-set increases the utilization of train-set. As long as the requirements of connection time are met, a train-set runs under a number of lines to operate, which will enhance the flexibility of operating train-set.

However, train-set running in uncertain operating sections will also bring a lot of negative factors, such as the difficulties in scheduling train-set circulation and rescheduling when disruption arises.



(a) Fixed use of train-set: 4 train-sets are required



(b) Flexible use of train-set: 3 train-sets are required

Figure 5.20: Experiments using different strategies of train-sets

5.5.3 Summary of the Computational Results

In this section, we report on the performance of the models and the solution techniques for integrating train-sets circulation into the ATP problem from Chapter 4.

The numerical investigation consisted of four parts. In Part 1 different control parameters for tolerance of disruption, composed of allowable adjustment and periodic structure, are investigated. In Part 2 the objective that only considers minimizing the required train-set is tested and compared with part 1. In Part 3 time window constraints are used to enforce adherence to the additional schedule. In Part 4 two different strategies of train-set application are compared.

From numerical investigations it is observed that the settings of perturbation tolerance, various objectives, the setting of time window and the use of train-set will effect the inserting solution and the number of require train-set.

The choice of higher tolerance, including larger allowed adjustments and deviations to the periodic structure, has an obvious effect on decreasing the number of required train-sets. It may however take quite some time to solve the problem, but we argued that setting an appropriate relative gap can provide an effective stopping criterion. The computational time to get a “nearly optimal” solution yields a substantial decrease. The result in Part 1 is further compared with that in Part 2 which applies the single objective function only to minimize the number of train-sets. The same or less train-sets are required and the computational time is decreased dramatically in Part 2. Building the additional schedule from scratch using time windows is effective, since both the search space for insertion and the potential option for train-set route are decreased dramatically. The ATP model can be applied both with the fixed and flexible use of train-set. The flexible use of train-set enhances the flexibility of operating train-set and less train-sets are required by adopting this approach.

Chapter 6

Conclusions and Future Work

The problem of inserting additional train services in an existing cyclic timetable is considered in this thesis. The primary motivation of the research occurs as a result of current operational problems in China's HSR. The technique of inserting new trains can be applied both in tactical planning that focus on the construction of a generic "cyclic + non-cyclic" timetable and short-term planning that concerns the re-development of a generic timetable in order to adapt to the demands of the individual weeks or days.

This concluding chapter first presents an overview of the main results of the thesis. We next discuss how these results provide answers to the research questions posed in Chapter 1. Finally, we reflect on the limitations of the thesis, and propose some directions for further research on inserting extra train paths in an existing timetable.

6.1 Main Results

In discussing the main results of the thesis, we distinguish between results regarding the applicability analysis of "cyclic + non-cyclic" timetable, the modelling of adding train paths (ATP) problem, the integration of train-set circulation to the ATP problem and the development of an aid tool for ATP problem.

"Cyclic + non-cyclic" Timetable

Having sketched the organizational environment for constructing railway timetables in China's HSR in Chapter 1, Chapter 2 subsequently analyses the appropriate mode of timetable for China's HSR. High quality of systems and infrastructures make China's HSR possible to operate trains cyclically to improve the quality of service. However, due to too many train ODs and some special trains, such as night train, long-distance and cross-line trains, an incomplete cyclic train operation mode is more practical. In particular, the "cyclic + non-cyclic" mode of timetable is proposed in this thesis, and its applicability and the planning process are discussed in detail. As such, we believe to have provided a valuable contribution, since most literature describes a hybrid timetable

only briefly, and none of them goes into details.

Modeling Adding Train Paths Problem

Having analysed the relations and differences among timetable scheduling, rescheduling and the ATP problem in Chapter 1, we define the ATP problem in Chapter 3 by stating our assumptions, requirements and considered objectives.

The ATP problem is different from the usual timetable construction problem due to additional constraints of tolerance of disruption to initial trains. These tolerance constraints may be viewed as the allowable adjustments and periodic structure. In this paper both settings are provided for in our techniques and investigated in our numerical investigations. Chapter 4 firstly formulate a basis model for the ATP problem in Section 4.1. In this work, the problem is characterized based on event-activity graph. The general constraints, such as flexible running times, dwell times, headways and time windows are modelled in the basic ATP model. Several extensions to the basic model are further proposed in Section 4.2. The real-world constraints that concerning the acceleration and deceleration times, priority for overtaking, station capacity, allowed adjustments, periodic structure and frequency of services are incorporated into the model. In order to get a new timetable that with quality of the performance to the additional trains, low deviations to the initial services and high quality of the entire trains, Chapter 4 presents the linear objective function of minimizing travel time of additional trains, minimizing total adjustments to initial trains, minimizing the makespan and maximizing the robustness of the new timetable. Later, in Chapter 5, we also introduce an objective function for minimizing the required number of train-sets.

The experimental studies consist of two parts, the initial timetable is fixed or unfixed. From the computational tests in Chapter 4, we conclude that the higher tolerance of frequency constraint is quite useful for handling a better robustness insertion. The computational tests further show that a balance objective (or multi-objective) function yield a substantial increase in the quality of the obtained insertions when the initial timetable can be changed, compared to the models that only consider the minimum trip time or maximum robustness.

Integration of train-set circulation and the ATP problem

In practice, the number of required train-sets should be also taken into account as an important index in order to obtain a match between the requested additional trains and the available number of train-sets. In Chapter 5, we integrate train-set circulation to the ATP problem from Chapter 4.

The train-set circulation in the ATP problem is decomposed to two sub-problem. For current train-set circulation, the initial train-set route is assumed to be fixed. For additional trains, different inserting patterns produce different train-set circulation and

number of required train-set consequently. The first sub-problem can be simply dealt with as a rescheduling problem of a tight constraint to keep the current train-set circulation. The second is a train-set routing problem to cover all the additional trains with minimum number of train-sets. The difficulty of the second sub-problem is that train-sets circulation are usually determined in the tactical planning phase after all of the train lines and timetable have been fixed. However, in the ATP problem, the additional trains do not exist in the initial timetable, and even the number of additional trains depends on the number of instantly available train-set at the right place. In order to solve the problem in a reasonable time, we start from fixed train-set route, and then apply flexible train-set route that provides possible alternative turning activities to decrease the waiting time of a train-set in an overnight turn-around. Combined with the ATP problem in Chapter 4, the model in Chapter 5 concerns to minimize the total adjustments for initial trains and at the same time minimize the number of required train-sets for entire trains.

The numerical investigation consists of four parts in Chapter 5. Several affecting factors are evaluated respectively, such as the various level of tolerance of disruptions, different objective functions, the introduction of time window constraints and the different use of train-sets. Within an hour of computation time, the proposed approach yields quite good insertions for our test instances.

Adding Train Paths Tool

A helpful tool for the ATP problem is developed based on the proposed models and approaches for testing the all of the experimental scenarios. This tool is programmed in Visual Basic language, as shown in Appendix C.

6.2 Answering the Research Questions

Having summarized the results of the thesis, we now return to the research questions defined in Chapter 1.

Why is the “cyclic + non-cyclic” timetable mode more appropriate in China’s HSR, and what is the process of planning such a hybrid timetable ?

As an answer to this question, we distinguish the operating characteristics between the China’s HSR, the conventional railway and the HSR in other countries in Chapter 2. By comparing the operating conditions in China’s HSR to the cyclic timetable experiments around the world, the applicability analysis and planning process of “cyclic + non-cyclic” timetable is decied in Section 2.3.

What real-word requirements should be taken into account, what are the criteria for assessing the quality of an insertion, and how can they be modelled?

Several real-world constraints for inserting additional trains are discussed in Section 3.3. The deviation to the initial cyclic timetable is limited, especially the periodic structure which is preferred to be kept to the most possible extent. The flexible trip time is also considered in this problem to enlarge solution space; the acceleration and deceleration time are added into the trip time when train stops. Priority constraints provide an criteria that which train should wait when track capacity is violated. The frequency of additional services is also constrained; we further avoid to fix a timetable beforehand by introducing a bandwidth. Section 4.1.3 and Section 4.2 describe how to model these constraints based on an event-activity graph.

Several criteria for evaluating the quality of an insertion are discussed in Section 3.2. We distinguish between the criteria of deviations to initila trains, travel time, timetable robustness, and makespan cost control. We argue that these criteria may be conflicting, and delineate each of them with respect to our purposes and assumptions. Section 4.3 describes how to model these criteria as objective functions in our ATP problem.

What adjustments need to be made to the adding train paths model to integrate the train-set circulation, and how can these adjustments be modelled?

For the sake of avoiding large disruptions to the scheduled services and solving adding paths problem within a reasonable computational force, the train-set planning in the ATP problem is decomposed into two sub-problems. For the initial timetable, the current train-set circulation is assumed to be fixed beforehand. It is solved as an rescheduling problem with a tight constraint, that a train-set operates the existing trains in the same sequence as it is scheduled in the initial timetable. For the additional trains, the train-set circulation problem is equivalent to covering all the additional trains with minimal number of train-sets. It is formulated based on the concept of rotation which is widely used in the airline industry. Section 5.4 describes how to integrate the train-set circulation to the ATP problem.

How can the models arising from the previous two questions be solved in a reasonable amount of time?

We discuss the practical requirements of inserting additional trains in Section 3.3. Satisfying all of these real-world constraints may firstly yield a substantial decrease on the computational time. From numerical investigations in Section 4.4 and Section 5.5, it is observed that the limited disruption to the initial trains, the introduction of time window and the frequency constraint will narrow the solution space. We further find that the computation times can be reduced by setting an appropriate relative gap which may provide an effective stopping criterion to a “nearly optimal” solution.

Moreover, in the ATP problem, the objective of minimum train-set is nonlinear since the additional trains do not exist in the timetable. In order to linearize the model and solve the problem in an reasonable time, we start from fixed train-set route, and then

apply flexible train-set route. In the latter process, we deal with the overnight turn-arounds and provide them possible alternative turning activities to decrease the waiting time of a train-set. Section 5.4 describes and formulates this process.

6.3 Limitations of the Thesis and Recommendations for Further Research

Obviously there are some aspects disregarded in the current formulation and due to the chosen level of detail and which may affect the quality and practical relevance of generated solutions. The switches especially in through station, for example, have not been modelled explicitly. That is, unless the turn around activity with a direction change and another trip are separated in time, their paths are not considered to be in conflict, while in practise they may be. An example would be that the turn around activity from train t_1 to t_3 is permitted by the formulation through the route r_5 while another train simultaneously uses route r_6 or r_7 . In practise, that would constitute a violation of the safety restrictions. Thus, additional research on this topic is required but beyond the scope of this paper.

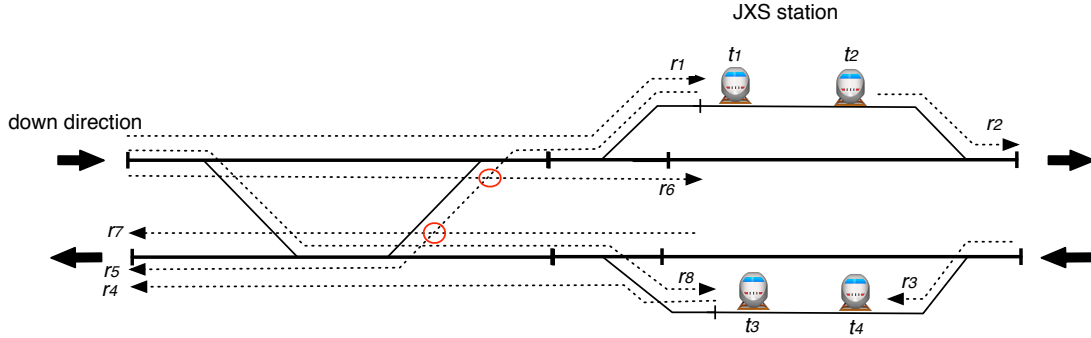


Figure 6.1: Turn around activity in a through station

Moreover, we have used standard software, CPLEX, with its branch-and-bound solution procedure and default settings of parameters. There may be more beneficial settings than the default settings (including branching strategies) for this particular problem. Hence, using more tailored solution software or parameter settings could potentially provide good solutions faster (Törnquist and Persson (2007)). Furthermore, ideally the inserting of traffic should be carried out with a network perspective and in a whole day time horizon, but the problem would become too large to solve within a reasonable time. Consequently the problem needs to be bounded somehow both in time and geographically. However, costs and gains that arise beyond the problem boundary should somehow be approximated and accounted for when considering a fragment of the overall inserting problem (Törnquist and Persson (2007)). In ongoing and future research, the

development of algorithms, able to find near optimal solutions for large instances within acceptable computation time is worthwhile.

Another practical consideration that has to be taken into account in the future is the consideration of multiple objectives. We have seen several evaluation criteria of practical relevance for an insertion, including minimization of adjustments, minimization of travel time, maximization of robustness, minimization of makspan and minimization of required train-set.

Our model considers the multi-objective combined models with an objective function being the weighted sum of the original objective functions. How to define and use suitable parameters in the objective function to represent the trade-off between various criteria still need to be discussed in detail. In addition, all of the initial trains have the same value of penalty to be adjusted in this paper. Applying various penalties to high-speed and middle-speed trains for example may lead to middle-speed trains becoming less prioritised than high-speed trains. Furthermore, other principle approaches for considering multiple objectives simultaneously, such as Pareto optimal solutions, may be considered.

Bibliography

- Abbas-Turki, A., Zaremba, E., Grunder, O., El-moudni, A., 2011. Perfect homogeneous rail traffic: A quick efficient genetic algorithm for high frequency train timetabling, in: Intelligent Transportation Systems (ITSC), 2011 14th International IEEE Conference on, pp. 1495–1500.
- Acuna-Agost, R., Michelon, P., Feillet, D., Gueye, S., 2010. A MIP-based local search method for the railway rescheduling problem. *Networks* 57, 69–86.
- Acuna-Agost, R., Michelon, P., Feillet, D., Gueye, S., 2011. SAPI: Statistical Analysis of Propagation of Incidents. A new approach for rescheduling trains after disruptions. *European Journal of Operational Research* 215, 227–243.
- Alfieri, A., Groot, R., Kroon, L., Schrijver, L., 2002. Efficient Circulation of Railway Rolling Stock. *ERIM Report Series Research in Management* , 1–23.
- Anderegg, L., Eidenbenz, S., Gantenbein, M., Stamm, C., Taylor, D.S., Weber, B., Widmayer, P., 2002. Train routing algorithms: concepts, design choices and practical considerations, in: Proceedings of the fifth workshop on algorithm engineering and experiments (ALENEX03), pp. 1–10.
- Ben-Khedher, N., Kintanar, J., Queille, C., Stripling, W., 1998. Schedule optimization at sncf: From conception to day of departure. *Interfaces* 28, 6–23. doi:10.1287/inte.28.1.6.
- Budai, G., Maróti, G., Dekker, R., Huisman, D., Kroon, L., 2009. Rescheduling in passenger railways: the rolling stock rebalancing problem. *Journal of Scheduling* 13, 281–297.
- Bunte, S., Kliwer, N., 2010. An overview on vehicle scheduling models. *Public Transport* 1, 299–317.
- Burdett, R., Kozan, E., 2009. Techniques for inserting additional trains into existing timetables. *Transportation Research Part B: Methodological* 43, 821–836.
- Cacchiani, V., Caprara, A., Toth, P., 2010. Scheduling extra freight trains on railway networks. *Transportation Research Part B: Methodological* 44, 215–231.

- Cacchiani, V., Huisman, D., Kidd, M., Kroon, L., Toth, P., Veelenturf, L., Wagenaar, J., 2014. An overview of recovery models and algorithms for real-time railway rescheduling. *Transportation Research Part B* 63, 15–37.
- Cacchiani, V., Toth, P., 2012. Nominal and robust train timetabling problems. *European Journal of Operational Research* 219, 727–737.
- Cadarso, L., Marín, Á., 2011. Integration of timetable planning and rolling stock in rapid transit networks. *Annals of Operations Research* 199, 113–135.
- Campos, J., de Rus, G., 2009. Some stylized facts about high-speed rail: A review of HSR experiences around the world. *Transport Policy* 16, 19–28.
- Castillo, E., Gallego, I., Ureña, J.M., Coronado, J.M., 2011. Timetabling optimization of a mixed double- and single-tracked railway network. *Applied Mathematical Modelling* 35, 859–878.
- Central Japan Railway Company, 2013. Central Japan Railway Company Annual Report 2013. Technical Report.
- Chung, J., Oh, S., Choi, I., 2009. A hybrid genetic algorithm for train sequencing in the Korean railway. *Omega* 37, 555–565.
- Cordeau, J.F., Desautniers, G., Lingaya, N., Soumis, F., Desrosiers, J., 2001. Simultaneous locomotive and car assignment at VIA Rail Canada. *Transportation Research Part B: Methodological* 35, 767–787.
- Corman, F., D’Ariano, A., Hansen, I., 2011. Effectiveness of dynamic reordering and rerouting of trains in a complicated and densely occupied station area. *Transportation Planning and Technology, Special Issue: Railway Operations Modelling and Analysis* 34, 341–362.
- Corman, F., D’Ariano, A., Hansen, I.A., 2010a. Disruption Handling In Large Railway Networks, in: *Computers in Railways XII*, WIT Press, Beijing, China. pp. 629–640.
- Corman, F., D’Ariano, A., Pacciarelli, D., Pranzo, M., 2009. Evaluation of green wave policy in real-time railway traffic management. *Transportation Research Part C* 17, 607–616.
- Corman, F., D’Ariano, A., Pacciarelli, D., Pranzo, M., 2010b. A tabu search algorithm for rerouting trains during rail operations. *Transportation Research Part B* 44, 175–192.
- Corman, F., D’Ariano, A., Pacciarelli, D., Pranzo, M., 2010c. Centralized Versus Distributed Systems to Reschedule Trains in Two dispatching areas. *Public Transport* 2, 219–247.

- Corman, F., D'Ariano, A., Pacciarelli, D., Pranzo, M., 2012. Bi-objective conflict detection and resolution in railway traffic management. *Transportation Research Part C* 20, 79–94.
- D'Ariano, A., Pacciarelli, D., Pranzo, M., 2007a. A branch and bound algorithm for scheduling trains in a railway network. *European Journal of Operational Research* 183, 643–657.
- D'Ariano, A., Pacciarelli, D., Pranzo, M., 2008a. Assessment of flexible timetables in real-time traffic management of a railway bottleneck. *Transportation Research Part C: Emerging Technologies* 16, 232–245.
- D'Ariano, A., Pacciarelli, D., Pranzo, M., 2008b. Reordering and Local Rerouting Strategies to Manage Train Traffic in Real Time. *Transportation Science* 42, 405–419.
- D'Ariano, A., Pranzo, M., 2008. An Advanced Real-Time Train Dispatching System for Minimizing the Propagation of Delays in a Dispatching Area Under Severe Disturbances. *Networks and Spatial Economics* 9, 63–84.
- D'Ariano, A., Pranzo, M., Hansen, I.A., 2007b. Conflict Resolution and Train Speed Coordination for Solving Real-Time Timetable Perturbations. *IEEE Transactions on Intelligent Transportation Systems* 8, 208–222.
- Dollevoet, T., 2013. Delay Management and Dispatching in Railways . Ph.D. thesis. Erasmus University Rotterdam.
- Dollevoet, T., Huisman, D., Kroon, L., Schmidt, M., Schbel, A., 2014. Delay management including capacities of stations. *Transportation Science* 0, null. doi:10.1287/trsc.2013.0506.
- Dollevoet, T., Huisman, D., Schmidt, M., Schöbel, A., 2012. Delay Management with Rerouting of Passengers . *Transportation Science* 46, 74–89.
- Feigenbaum, B., 2013. High-Speed Rail in Europe and Asia: Lessons for the United States. URL: http://reason.org/files/high_speed_rail_lessons.pdf.
- Fioole, P.J., Kroon, L., Maróti, G., Schrijver, A., 2006. A rolling stock circulation model for combining and splitting of passenger trains. *European Journal of Operational Research* 174, 1281–1297.
- Fischetti, M., Salvagnin, D., Zanette, A., 2009. Fast approaches to improve the robustness of a railway timetable. *Transportation Science* 43, 321–335. doi:10.1287/trsc.1090.0264.
- Flier, H., Graffagnino, T., Nunkesser, M., 2009. Scheduling Additional Trains on Dense Corridors. *Lecture Notes in Computer Science* 5526/2009, 149–160.

- Flier, H., Nunkesser, M., Schachtebeck, M., Schöbel, 2008. Integration Rolling Stock Circulation into the Delay Management Problem. Project ARRIVAL Number FP6-0212235-2 .
- Geraets, F., Kroon, L., Schöbel, A., Wagner, D., D Zaroliagis, C., 2004. Algorithmic Methods for Railway Optimization. International Dagstuhl Workshop , 1–333.
- Ghoseiri, K., Szidarovszky, F., Asgharpour, M.J., 2004. A multi-objective train scheduling model and solution. *Transportation Research Part B: Methodological* 38, 927–952.
- Hansen, I.A., 2009. Railway Network Timetabling and Dynamic Traffic Management, in: 2nd International Conference on Recent Advances in Railway Engineering (ICRARE-2009), Iran university of science and Technology Tehran I.R. Iran. pp. 1–11.
- Hansen, I.A., Pachl, J., 2008. *Railway Timetable & Traffic*. Eurailpress —DVV Rail Media.
- Hong, S., Kim, K., Lee, K., Hwanpark, B., 2009. A pragmatic algorithm for the train-set routing: The case of Korea high-speed railway. *Omega* 37, 637–645.
- Ingolotti, L., Barber, F., Tormos, P., Lova, A., Salido, M., Abril, M., 2004. An Efficient Method to Schedule New Trains on a Heavily Loaded Railway Network. *Lecture Notes in Computer Science* 3315/2004, 164–173.
- Jia, X., 2011. A Study on Optimization Theories and Methods about The Periodic Train Work Diagram of the Passenger Dedicated Line. Ph.D. thesis. Southwest Jiaotong University.
- Khan, M.B., 2008. Slack time allocation in robust double-tracked railway timetable optimization. Ph.D. thesis. Southwest Jiaotong University.
- KPMG Global, C., 2013. China 360 All aboard: High-speed rail network connecting China. Technical Report.
- Leon, P., 2003. cyclic railway timetable optimization. Ph.D. thesis. Erasmus Universiteit Rotterdam.
- Li, H.C., Liu, Z.N., Yang, Y.P., 2011. Inspection report: Some suggestions to railway industry of China and transportation development urban railway transit of Beijing city. *Railway Economic Research* 3, 17–26.
- Lindner, T., 2000. Train Schedule Optimization in Public Rail Transport. Ph.D. thesis. Technische Universiteit Braunschweig.
- Lingaya, N., Cordeau, J.F., Desaulniers, G., Desrosiers, J., Soumis, F., 2002. Operational car assignment at VIA Rail Canada. *Transportation Research Part B: Methodological* 36, 755–778.

- Lloyd, C., Ellis, J., George, N., Zhu, Z., 1997. The aircraft rotation problem. *Annals of Operations Research* 69, 33–46.
- Lusby, R., Larsen, J., Ehrgott, M., Ryan, D., 2011a. Railway track allocation: models and methods. *OR Spectrum* 33, 843–883. doi:10.1007/s00291-009-0189-0.
- Lusby, R., Larsen, J., Ryan, D., Ehrgott, M., 2011b. Routing trains through railway junctions: A new set-packing approach. *Transportation Science* 45, 228–245. doi:10.1287/trsc.1100.0362.
- Mackenzie, S., 2000. Train timetabling on complex networks, in: *Proceedings from the Conference on Railway Engineering 2000 (CORE2000)*, Adelaide, Australia.
- Maróti, G., 2006. *Operations Research Models for Railway Rolling Stock Planning*. Ph.D. thesis. Technische Universiteit Eindhoven. Technische Universiteit Eindhoven.
- Mascis, A., Pacciarelli, D., 2002. Job-shop scheduling with blocking and no-wait constraints. *European Journal of Operational Research* 143, 498 – 517. doi:http://dx.doi.org/10.1016/S0377-2217(01)00338-1.
- Meng, L., Zhou, X., 2014. Simultaneous train rerouting and rescheduling on an N-track network: A model reformulation with network-based cumulative flow variables. *Transportation Research Part B* 67, 208–234.
- Mu, S., Dessouky, M., 2011. Scheduling freight trains traveling on complex networks. *Transportation Research Part B* 45, 1103–1123.
- Nie, L., Fei, D.B., Zhou, S.D., Fu, H.L., Tong, L., 2010a. Key Issues on Train Line Planning for Beijing-Shanghai High Speed Railway, in: *Proceedings of the 2010 Joint Rail Conference*, pp. 1–6.
- Nie, L., Hansen, I.A., 2005. System analysis of train operations and track occupancy at railway stations. *European Journal of Transport and Infrastructure Research* 5, 31–54.
- Nie, L., HU, B., Fu, H., TONG, L., 2010b. Interaction Analysis between Night Train Operation and Maintenance Time on Passenger Dedicated Railway Line. *Journal of Transportation Systems Engineering and Information Technology* 10, 66–72.
- Nielsen, L.K., 2011. *Rolling Stock Rescheduling in Passenger Railways: Applications in short-term planning and in disruption management*. Ph.D. thesis. Erasmus University Rotterdam.
- Nielsen, L.K.r., Kroon, L., Maróti, G., 2012. A rolling horizon approach for disruption management of railway rolling stock. *European Journal of Operational Research* 220, 496–509.

- Oliveira, E., Smith, B.M., 2001. A hybrid constraint-based method for single-track railway scheduling problem. Technical Report.
- Pacciarelli, D., Pranzo, M., 2001. A tabu search algorithm for the railway scheduling problem.
- Pachl, J., 2009. Railway Operation and Control, second edition. VTD Rail Publishing.
- Peeters, M., Kroon, L., 2008. Circulation of railway rolling stock: a branch-and-price approach. *Computers & Operations Research* 35, 538–556.
- Pisinger, D., Ropke, S., 2007. A general heuristic for vehicle routing problems. *Computers & Operations Research* 34, 2403–2435.
- Potthoff, G., 1980. Verkehrsströmungslehre, Band1 - Die Zugfolge auf Strecken und in Bahnhöfen. 3rd Edition. transpress Verlag.
- Rouillon, S., Desaulniers, G., Soumis, F., 2006. An extended branch-and-bound method for locomotive assignment. *Transportation Research Part B: Methodological* 40, 404–423.
- Schachtebeck, M., 2010. Delay Management in Public Transportation: Capacities, Robustness, and Integration. Ph.D. thesis. Georg-August-Universität Göttingen.
- Schachtebeck, M., Schöbel, A., 2010. To wait or not to wait and who goes first? Delay Management with Priority Decisions. *Transportation Science* 44, 307–321.
- Schmidt, M., 2013. Simultaneous optimization of delay management decisions and passenger routes. *Public Transport* 5, 125–147.
- Schöbel, A., 2001a. Integer programming approaches for solving the delay management problem. *Algorithmic Methods for Railway Optimization*, 145–170.
- Schöbel, A., 2001b. A model for the delay management problem based on mixed-integer-programming. *Electronic Notes in Theoretical Computer Science* 50, 1 – 10. doi:[http://dx.doi.org/10.1016/S1571-0661\(04\)00160-4](http://dx.doi.org/10.1016/S1571-0661(04)00160-4). {ATMOS} 2001, Algorithmic Methods and Models for Optimization of RailwayS (Satellite Workshop of {ICALP} 2001).
- Schöbel, A., 2007. Integer programming approaches for solving the delay management problem, in: Geraets, F., Kroon, L., Schoebel, A., Wagner, D., Zaroliagis, C. (Eds.), *Algorithmic Methods for Railway Optimization*. Springer Berlin Heidelberg. volume 4359 of *Lecture Notes in Computer Science*, pp. 145–170. doi:10.1007/978-3-540-74247-0_7.

- Schöbel, A., 2009. Capacity constraints in delay management. *Public Transport* 1, 135–154.
- Serafini, P., Ukovich, W., 1989. A mathematical model for periodic scheduling problems. *SIAM Journal on Discrete Mathematics* 2, 550–581. doi:10.1137/0402049.
- Shafia, M.A., Pourseyed Aghaee, M., Sadjadi, S.J., Jamili, A., 2012. Robust Train Timetabling Problem: Mathematical Model and Branch and Bound Algorithm. *IEEE Transactions on Intelligent Transportation Systems* 13, 307–317.
- Sun, Q., Feng, X., Bian, K., 2011. Operation and Organization Management of High-speed Railway in Japan. *Journal of Transportation Systems Engineering and Information Technology* 11, 11–16.
- Tan, Y., 2013. A bound search algorithm and iterative refinements for scheduling extra train paths. *TRB 92nd Annual Meeting Compendium of Papers* .
- Thompson, L.S., Tanaka, Y., 2011. High Speed Rail Passenger Services: World Experience and U.S. Applications. Technical Report.
- Törnquist, J., 2006. Computer-based Decision Support for Railway Traffic Scheduling and Dispatching: A Review of Models and Algorithms. *DROPS-IDN/659* .
- Törnquist, J., 2012. Design of an effective algorithm for fast response to the re-scheduling of railway traffic during disturbances. *Transportation Research Part C* 20, 62–78.
- Törnquist, J., Persson, J.A., 2007. N-Tracked Railway Traffic Re-Scheduling During Disturbances. *Transportation Research Part B: Methodological* 41B.
- Vaidyanathan, B., Ahuja, R.K., Liu, J., Shughart, L.A., 2008. Real-life locomotive planning: New formulations and computational results. *Transportation Research Part B: Methodological* 42, 147–168.
- Wang, B., 2008. Researches of Line Planning and Timetable Scheduling Based on Periodic operations for the Rail Traffic in China . Ph.D. thesis. Beijing Jiaotong University.
- Wang, X., Huang, S., Zou, T., Yan, H., 2012. Effects of the high speed rail network on China's regional tourism development. *TMP* 1, 34–38.
- Xie, M., 2010. Railway Train-set Circulation Based on Timetable Optimization. Ph.D. thesis. Beijing Jiaotong University.
- Xu, X.F., Yang, X.J., 2011. Study on relative problems of transport organization on high speed railway. *Traffic and Transportation* 12, 84–87.
- Y, W., J, L., J, M.J., 2009. The discussion of organization in transportation about the intercity rail passenger line. *Railway Transport and Economy* 31, 20–22.

- Yang, D., Nie, L., Tan, Y., He, Z., Zhang, Y., 2010. Working out an incomplete cyclic train timetable for high-speed railways by computer. *WIT Transactions on the Built Environment* 114, 899–999.
- Zhou, X., Zhong, M., 2005. Bicriteria train scheduling for high-speed passenger railroad planning applications. *European Journal of Operational Research* 167, 752–771.
- Zhou, X., Zhong, M., 2007. Single-track train timetabling with guaranteed optimality: Branch-and-bound algorithms with enhanced lower bounds. *Transportation Research Part B: Methodological* 41, 320–341.
- Ziarati, K., Soumis, F., Desrosiers, J., Gelinas, S., Saintonge, A., 1997. Locomotive assignment with heterogeneous consists at CN North America. *European Journal of Operational Research* 97, 281–292.

Appendices

Appendix A

Different Mode of Maintenance Time

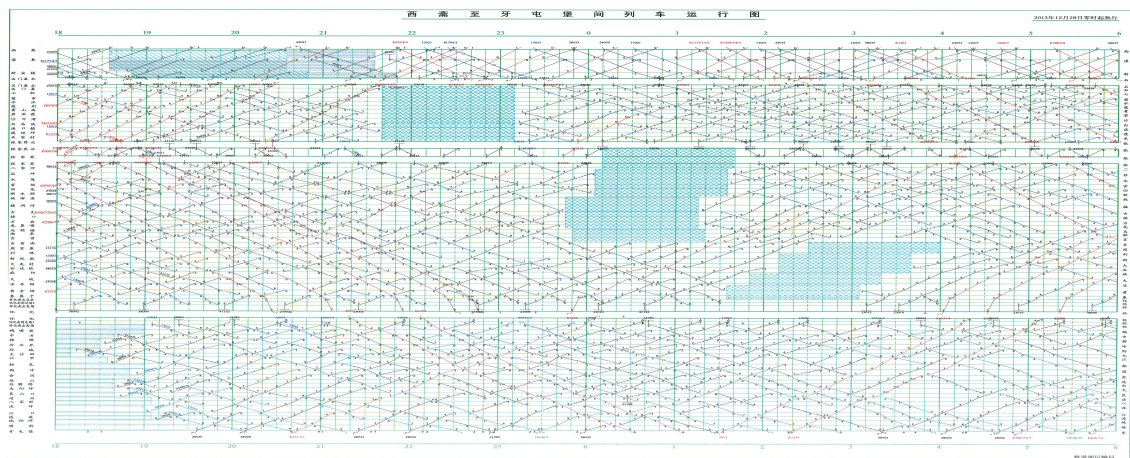


Figure A.1: Rectangular-shape maintenance time in conventional rail line

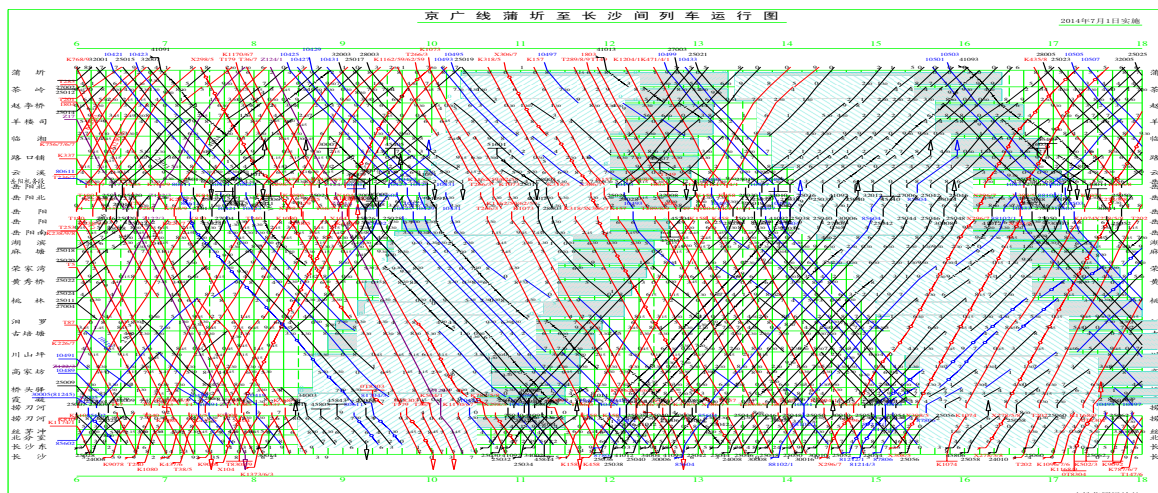


Figure A.2: V-shape maintenance time in conventional rail line

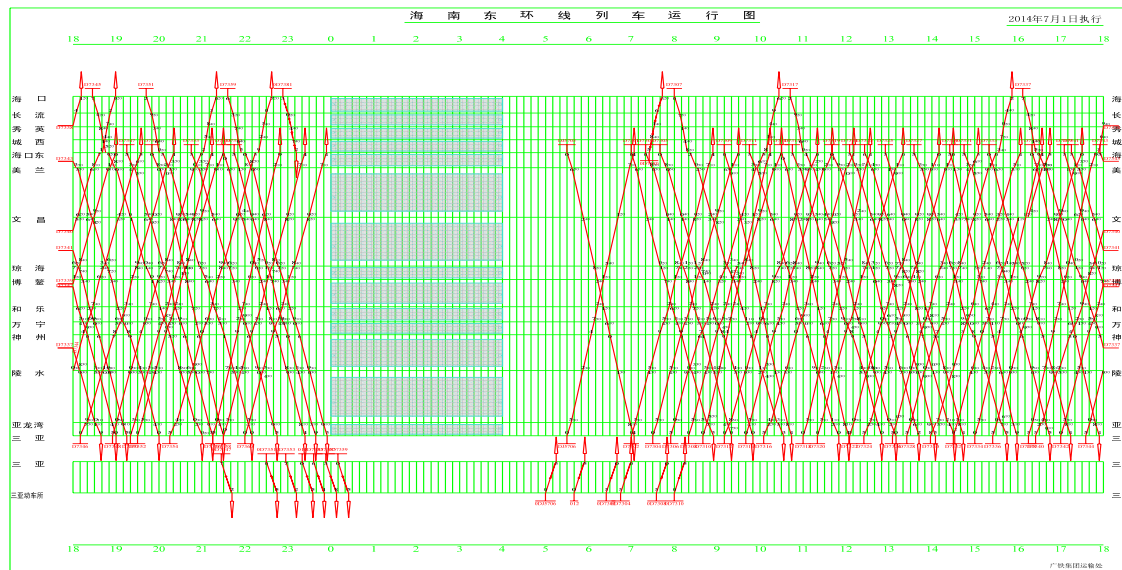


Figure A.3: Maintenance time in China's HSR line

Appendix B

The timetable in Guangzhou-Shenzhen HSR line

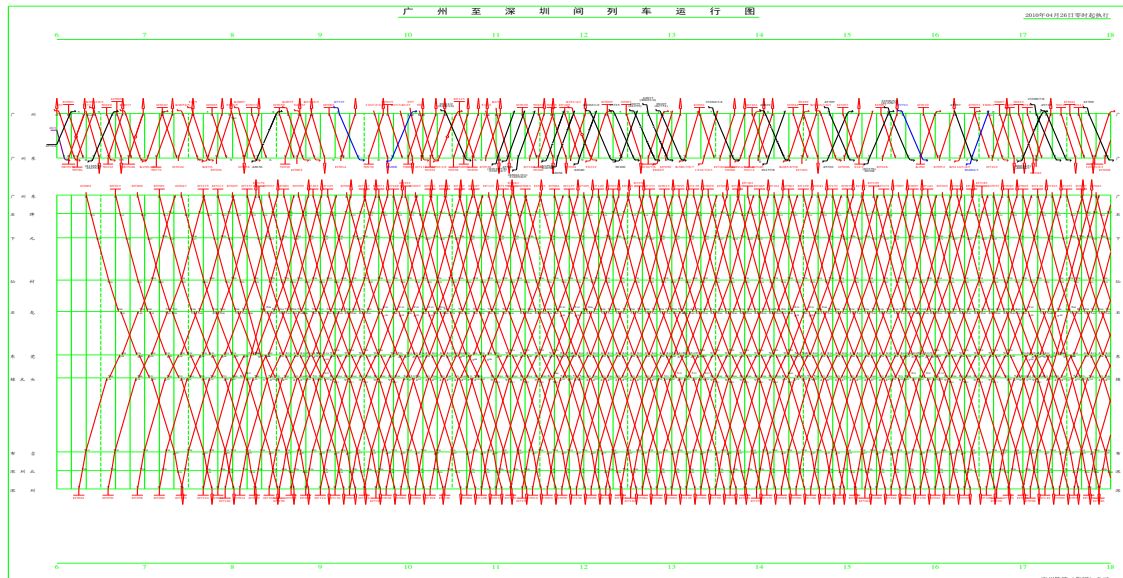


Figure B.1: The timetable in Guangzhou-Shenzhen HSR line

Appendix C

Graphical User Interfaces of the adding train paths tool

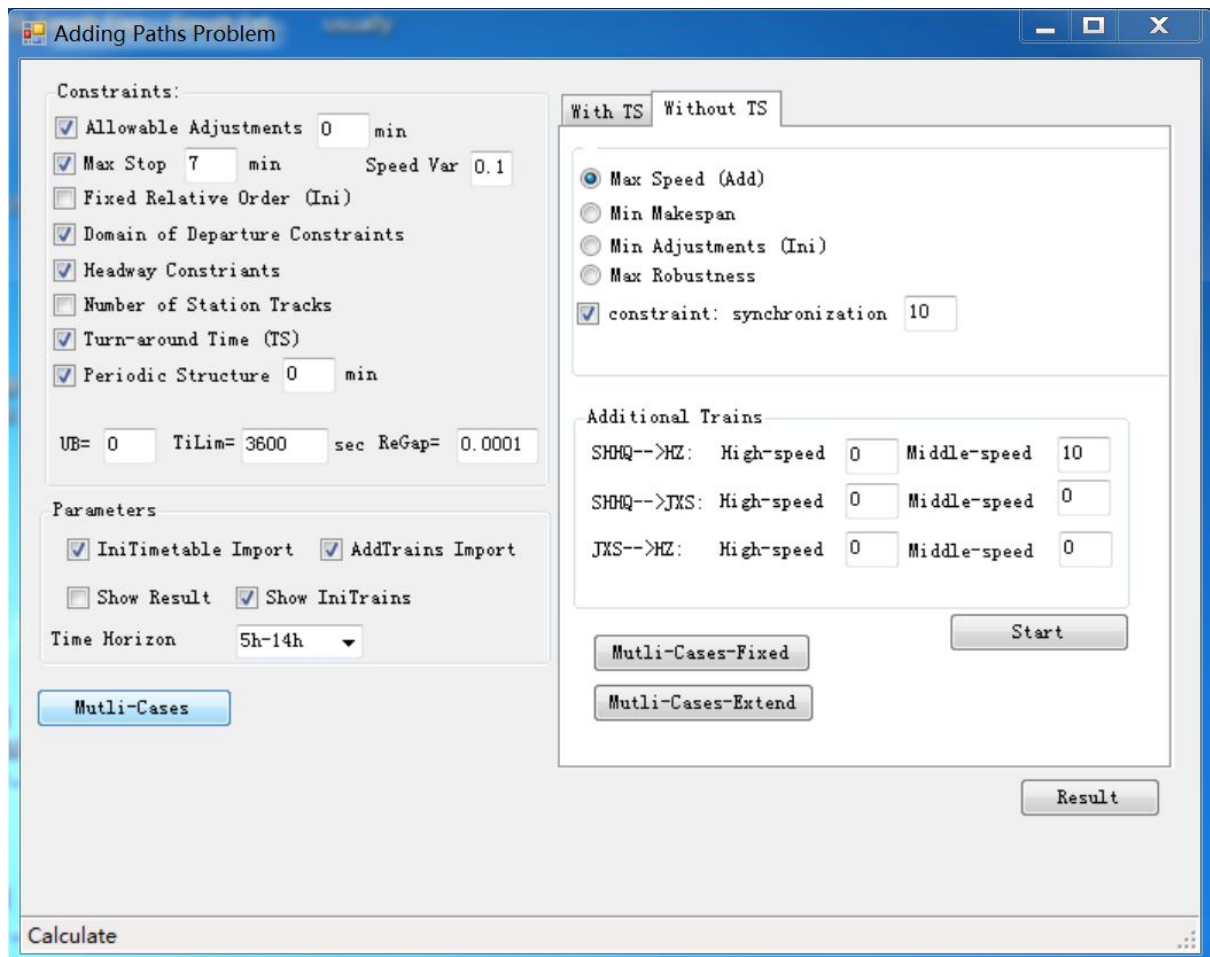


Figure C.1: The ATP problem without train-set circulation

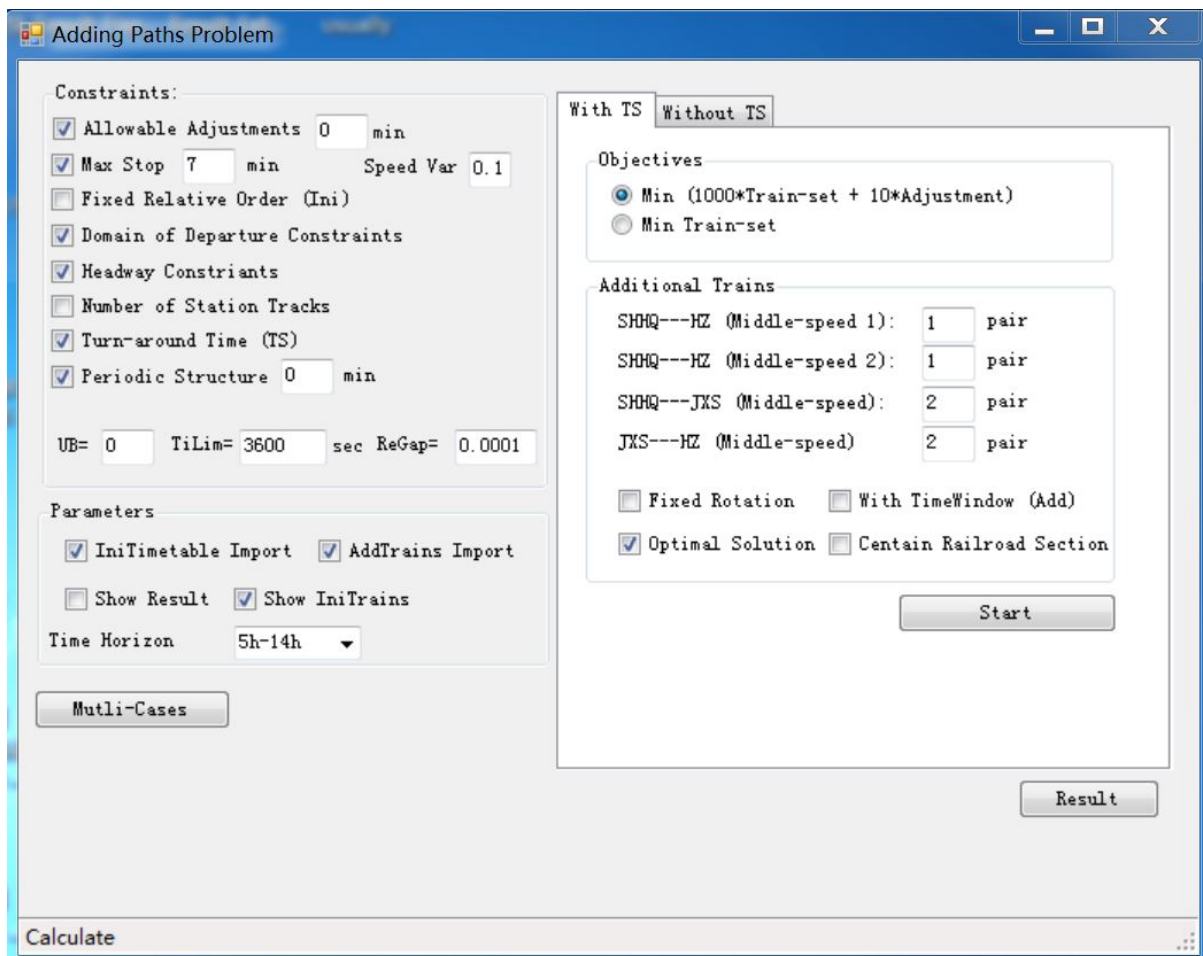


Figure C.2: The ATP problem with train-set circulation

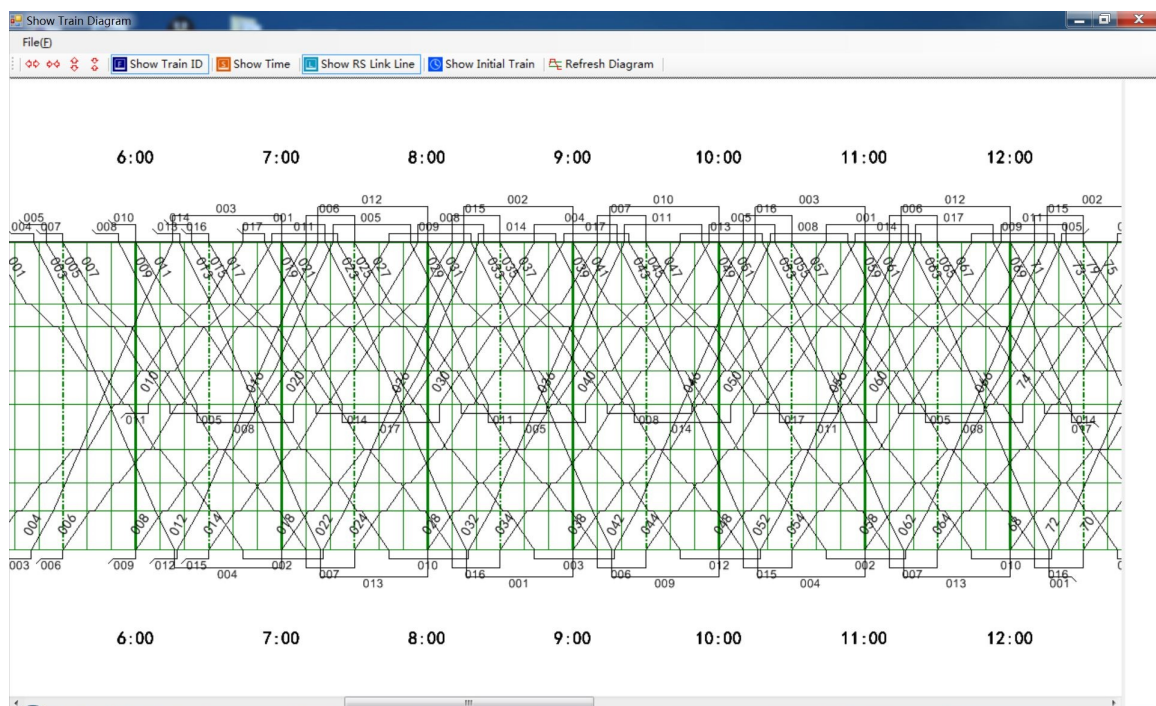


Figure C.3: Show result